



UNIVERSITÀ DI PISA

DIPARTIMENTO DI FARMACIA

Corso di Laurea Specialistica in Chimica e Tecnologia Farmaceutiche

Tesi di Laurea:

**STUDI COMPUTAZIONALI PER L'OTTIMIZZAZIONE DI NUOVI
INIBITORI DELLA FAAH**

Relatore: Prof. Tiziano Tuccinardi

Relatore: Dr. Giulio Poli

Candidato: Alessandro Villardita (**Matricola n° 457208**)

Settore Scientifico Disciplinare: **CHIM 08**

ANNO ACCADEMICO 2014-2015

INDICE

| | |
|--|-----------|
| CAPITOLO 1 INTRODUZIONE | 1 |
| 1.1 Vie Metaboliche degli Endocannabinoidi..... | 3 |
| 1.2 La FAAH | 5 |
| 1.2.1 Sito catalitico della FAAH | 6 |
| 1.2.2 Meccanismo d'azione della FAAH | 8 |
| 1.3 Il Sistema Endocannabinoide Come Target Terapeutico | 10 |
| 1.3.1 Trattamento del dolore e dell'infiammazione | 10 |
| 1.3.2 Trattamento dei disordini del Sistema Nervoso Centrale | 11 |
| 1.3.2.1 Trauma cranico | 12 |
| 1.3.2.2 Sclerosi Multipla..... | 13 |
| 1.3.2.3 Morbo di Parkinson | 13 |
| 1.3.2.4 Epilessia..... | 14 |
| 1.3.2.5 Schizofrenia | 14 |
| 1.3.2.6 Ansia e depressione | 15 |
| 1.3.2.7 Insonnia..... | 16 |
| 1.3.2.8 Nausea e vomito | 16 |
| 1.3.3 Trattamento di patologie cardiovascolari | 17 |
| 1.3.4 Trattamento dell'asma..... | 18 |
| 1.3.5 Trattamento del cancro..... | 18 |
| 1.4 Inibitori della FAAH presenti in letteratura..... | 20 |
| CAPITOLO 2 METODI | 29 |
| 2.1 Studi di Docking | 29 |
| 2.1.1 Gold..... | 29 |
| 2.1.2 Glide | 35 |
| 2.1.3 AutoDock | 37 |
| 2.3.3.1 Preparazione dei file di input | 37 |

| | |
|--|-----------|
| 2.3.3.2 Generazione della griglia | 38 |
| 2.3.3.3 Inizio del calcolo di docking..... | 39 |
| 2.3.3.4 Autodock da shell | 40 |
| 2.1.4 Vina | 42 |
| 2.1.5 Dock | 45 |
| 2.3.5.1 Preparazione dei file di input | 46 |
| 2.3.5.2 Generazione delle sfere e della griglia di scoring..... | 48 |
| 2.3.5.3 Inizio del calcolo di docking..... | 50 |
| 2.1.6 Fred 3.0..... | 54 |
| 2.2 Studi di Dinamica Molecolare | 61 |
| 2.2.1 Preparazione Step by Step dei File di Input per la Dinamica Molecolare | 61 |
| 2.2.2 Minimizzazione e Simulazioni di Dinamica | 65 |
| 2.2.3 Valutazione della Dinamica Molecolare | 68 |
| 2.2.3.1 Analisi RMSD | 68 |
| 2.2.3.2. Analisi MMPBSA..... | 70 |
| 2.2.3.3 Analisi di Nuovi Legami a Idrogeno | 74 |
| 2.3 Creazione di un File Average per Ulteriori Studi Computazionali | 74 |
| CAPITOLO 3 PARTE SPERIMENTALE | 77 |
| 3.1 Analisi delle Strutture Cristallografiche | 77 |
| 3.2 Studio Preliminare di Consensus Docking sulla Pose VS8 | 78 |
| 3.3 Analisi di Docking Approfondite | 81 |
| 3.3.1 Analisi dei risultati di docking | 82 |
| 3.4 Studi di Dinamica Molecolare | 84 |
| 3.4.1 Simulazioni di dinamica molecolare e analisi dei risultati..... | 84 |
| 3.5 Creazione di un Average e Studio di Self-Consensus-Docking | 91 |
| CAPITOLO 4 CONCLUSIONI | 93 |
| BIBLIOGRAFIA | 95 |

CAPITOLO 1 INTRODUZIONE

Come ampiamente dimostrato [1], la marijuana ed alcuni dei suoi costituenti cannabinoidi influenzano il SNC in modo dose – dipendente. La loro natura idrofobica ha portato a paragonare l'azione di queste sostanze ad agenti anestetici, un'azione quindi che non perturbi la membrana cellulare ma che agisca su specifici recettori. Negli anni '90, grazie a Matsuda e colleghi [2], fu identificata una sequenza di Dna codificante un recettore accoppiato a proteine-G che presto si rivelò essere un target endogeno per il maggior costituente psicoattivo della Cannabis Sativa L., il Δ^9 -tetraidrocannabinolo (Δ^9 -THC) [3]. Questo recettore fu chiamato “recettore CB1 dei cannabinoidi” e nel 1993 fu poi riportata la caratterizzazione molecolare di una seconda isoforma recettoriale, il recettore cannabinoide CB2. La loro identificazione costituì uno dei maggiori passi verso la rappresentazione di un nuovo sistema di trasmissione del segnale nominato “sistema endocannabinoide” (ECs). Queste scoperte condussero alla questione di quali fossero le molecole endogene in grado di riconoscere e/o attivare tali recettori. Tra gli anni '92-'97 si arrivò quindi alla scoperta dei cosiddetti “endocannabinoidi”.

Nel 1992, Devane et al. [4], attraverso uno screen per ligandi endogeni dei recettori cannabinoidi, identificarono un derivato dell'acido arachidonico, l'Arachidoniletanolamina, o Anandamide, il primo ligando endogeno per i recettori CB1. I ricercatori si accorsero che l'Anandamide andava ad inibire in modo competitivo e concentrazione dipendente, specifici siti di legame su cui facevano legame sonde di cannabinoidi radiomarcati. La sua struttura (Fig.1) fu determinata attraverso la spettrometria di massa e la risonanza magnetica nucleare.

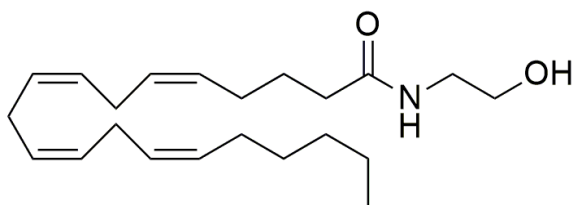


Fig.1 Struttura chimica Anandamide

Ad oggi l'Anandamide appare ancora come il principale rappresentante endogeno del sistema endocannabinoide, anche se omologhi e analoghi sono stati isolati o suggeriti. Si pensa agisca come un agonista parziale dei recettori dei cannabinoidi, senza alcuna significativa selettività fra il tipo CB1 o CB2. È ampiamente accettato che l'Anandamide possa interagire con ulteriori bersagli farmacologici, tra cui, ad esempio, recettori vanilloidi TRPV1 (transient receptor potential vanilloid type-1), i recettori nucleari PPAR (nuclear peroxisome proliferator-activated receptor) ed i recettori GPR55 (G-protein-coupled receptor 55) [5]. L'Anandamide, prodotta localmente, è stata vista giocare un ruolo fondamentale non solo nel SNC, dove modula diversi processi neurocomportamentali, come il dolore, l'alimentazione e l'ansia, ma anche nella maggior parte dei sistemi fisiologici che sono stati indagati, come il sistema immunitario, cardiovascolare, riproduttivo, respiratorio e scheletrico.

Nel 1995, venne identificato come secondo ligando cannabinoide endogeno, un monogliceride conosciuto, il 2-arachidonilglicerolo (2-AG) [6] [7]. Il 2-AG venne isolato dal tessuto intestinale e presenta una concentrazione, a livello cerebrale, 170 volte superiore rispetto all'Anandamide. Il 2-AG viene prodotto nell'ippocampo e la sua formazione è calcio dipendente, mediata dagli enzimi fosfolipasi C e diacilglicerolo lipasi. Contrariamente all'Anandamide il 2-AG attiva i recettori cannabinoidi a livello neuronale, come un'agonista totale, anch'esso senza distinzione per il sottotipo recettoriale, ed inibisce il potenziale d'azione a lungo termine [8].

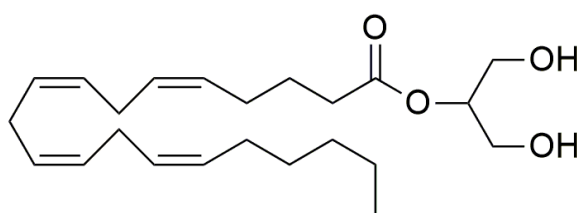


Fig.2 Struttura chimica 2-AG

1.1 Vie Metaboliche degli Endocannabinoidi

Una comprensione delle vie metaboliche degli endocannabinoidi è fondamentale per svelare il ruolo del sistema endocannabinoide in condizioni fisiologiche e patologiche e mettere a punto nuove strategie terapeutiche. L'attività dell'AEA e del 2-AG sui loro recettori viene conclusa rapidamente attraverso un processo a due step che inizia con il trasporto attraverso la membrana plasmatica, seguito da una reazione di idrolisi intracellulare catalizzata da differenti lipasi [9].

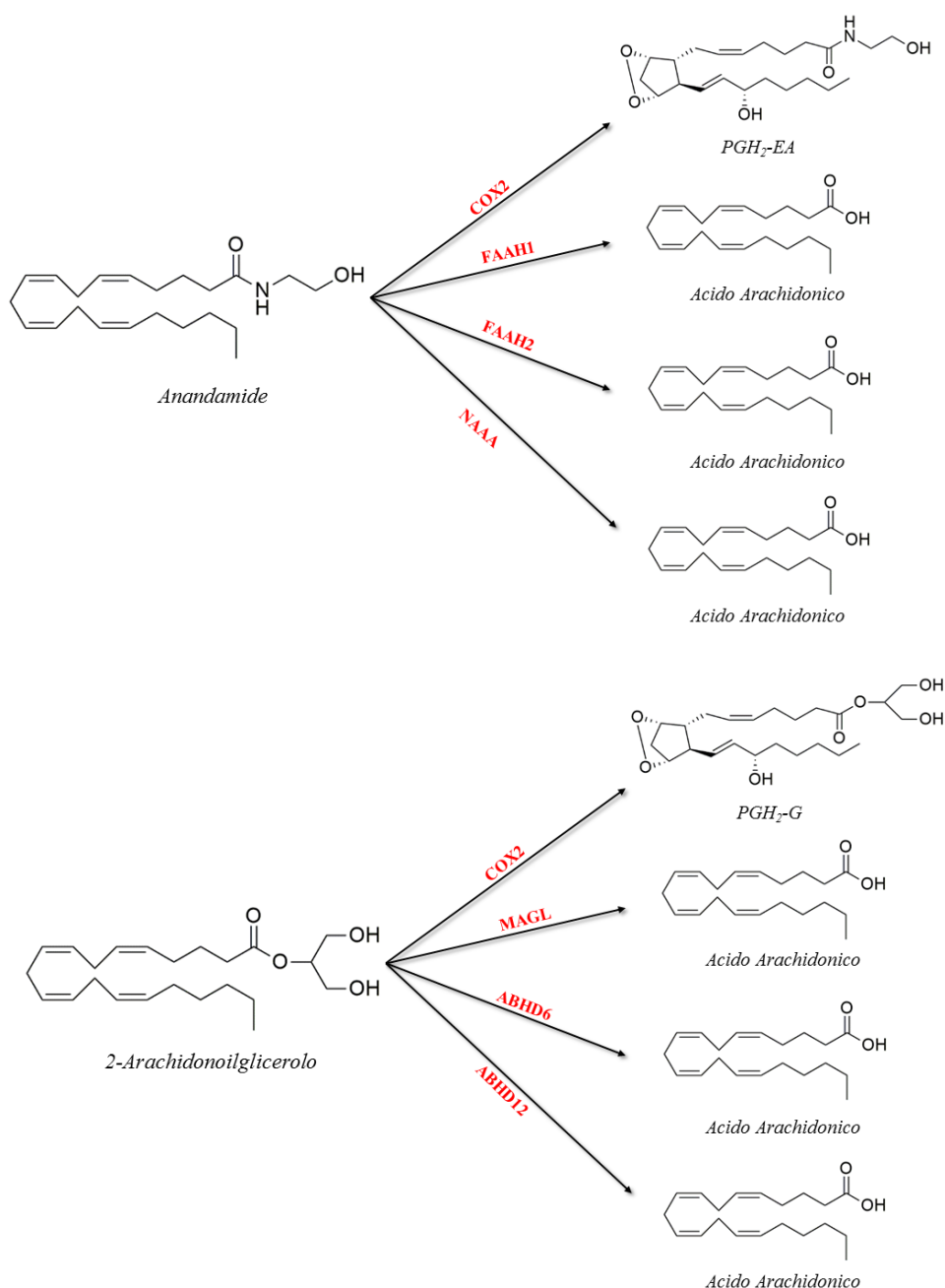


Fig.3 Vie metaboliche degli endocannabinoidi Anandamide e 2-AG

Le differenti lipasi coinvolte nella terminazione del segnale endocannabinoide, rappresentate in figura 3, sono:

COX (cyclooxygenase), in quanto derivati dell'acido arachidonico, sono anche oggetti alla maggior parte delle vie metaboliche ossidative che conducono alla biosintesi degli eicosanoidi [10];

FAAH1 (fatty acid amide hydrolase-1), la principale idrolasi deputata all'inattivazione dell'AEA;

NAAA (N-acyl ethanolamine acid amidase), altamente espressa nel sistema immunitario all'interno dei lisosomi, dove viene attivata mediante una scissione proteolitica [11];

FAAH-2 (fatty acid amide hydrolase-2), un isoforma della FAAH recentemente scoperta in tessuti umani ma non nei roditori. Sebbene mostri una triade catalitica identica alla FAAH, la localizzazione nell'organismo è diversa ed il livello di omologia tra le due isoforme è solamente del 20% [12];

MAGL (monoacylglycerol lipase), la principale idrolasi deputata all'inattivazione del 2-AG [14];

ABHD6 (α/β -hydrolase 6) e **ABHD12** (α/β -hydrolase 12) [13], recentemente scoperte nel SNC. Si pensa siano implicate nel metabolismo del 2-AG. La collocazione sub-cellulare differente rispetto alla MAGL suggerisce ruoli distinti nel controllare i livelli di 2-AG [13].

Sebbene gli endocannabinoidi siano lipofili in natura e possono quindi attraversare liberamente la membrana citoplasmatica, studi evidenziano il coinvolgimento di un meccanismo facilitato nell'assorbimento di AEA e 2-AG all'interno delle cellule, anche se la sua identità non è ancora stata divulgata e rimane materia di dibattito [14].

1.2 La FAAH

L'enzima FAAH (fatty acid amide hydrolase), isolato nel 1996 e ad oggi estensivamente caratterizzato, è stato identificato come la principale idrolasi responsabile della degradazione dell'Anandamide in vivo, sia a livello del SNC che periferico [15]. Essa è espressa in numerosi tessuti, quali cervello, intestino, fegato, testicoli, utero, rene, occhi, milza e polmone. Nel 2002, uno studio su topi "knock-out" per il gene della FAAH, in ugual modo a studi effettuati su animali trattati con inibitori covalenti della FAAH, hanno dimostrato come i livelli di Anandamide aumentino considerevolmente nel sistema nervoso centrale di tali soggetti. Inoltre questi roditori si sono mostrati più sensibili alla somministrazione di Anandamide ed hanno esibito un'ipoalgesia fenotipica [16].

L'attività della FAAH non è correlata solamente all'inibizione dell'anandamide, ma funge da regista anche in altri contesti farmacologici all'interno dell'organismo. In particolare la FAAH va ad inibire ligandi endogeni non cannabinoidi come l'N-palmitoiletanolamina (PEA), l'N-oleoiletanolamina (OEA), l'Oleamide, le N-acil taurine, associati rispettivamente ad attività antinfiammatoria, anoressigena, ipno-inducente ed analgesica [16]. Anche il 2-arachidonoilglicerolo è un ottimo substrato per la FAAH, ma questo solo in vitro poiché in vivo è stato dimostrato come né l'ablazione del gene della FAAH né l'inibizione farmacologica dell'enzima porti ad aumentare i livelli di 2-AG [17-18]

Affrontando l'evoluzione storica degli studi effettuati sulla FAAH, si può partire dal 1966 dove Bachur e Udenfriend [19] descrissero un'attività enzimatica in sistemi microsomiali, la quale catalizzava la formazione di ammidi degli acidi grassi, di etanolammine e di varie ammine farmacologicamente attive, e questo enzima mostrava una distribuzione tissutale simile alla FAAH. Negli anni '80 venne poi identificato un enzima localizzato all'interno dei mitocondri del fegato di ratto, chiamato ammido-idrolasi. Esso fu visto catalizzare l'idrolisi dell'N-aciletanolamina ad etanolamina e acidi grassi [20], ma un importante passo si ottenne nel 1996, quando Cravatt e collaboratori [21] riuscirono ad isolare la FAAH, originariamente chiamata oleamide idrolasi, dalla membrana plasmatica di epatociti di ratto.

1.2.1 Sito catalitico della FAAH

I numerosi sforzi fatti per ottenere strutture cristallografiche della FAAH umanizzata furono ostacolati dalle basse rese di espressione in sistemi ricombinanti e da proprietà biochimiche non eccelse, come l'instabilità e l'aggregazione, che rendevano difficile il processo di purificazione. Nel 2008 Mileni e collaboratori [22], riuscirono a descrivere una strategia diversa che prevedeva l'interconversione mutagenica dei siti attivi della FAAH di ratto (rFAAH) e della FAAH umana (hFAAH), ottenendo una FAAH di ratto "umanizzata" (r/h FAAH) che conteneva un sito attivo perfettamente identico a quello umano all'interno della struttura propria della proteina di ratto. Questa h/r FAAH esibiva lo stesso profilo di sensibilità inibitoria della h/FAAH, mantenendo al contempo le alte rese di espressione e le proprietà biochimiche dell'enzima di ratto.

La FAAH è una proteina integrale di membrana di 63kDa costituita da 579 amminoacidi e codificata da un gene localizzato sul cromosoma 1p34-35. Essa appartiene alla cosiddetta Amidase-Signature Family, i cui membri mostrano una sequenza amminoacidica comune che comprende circa 130 residui: "amidase-signature sequence". L' AS family è principalmente rappresentata nel regno dei funghi e dei batteri, e la FAAH è stata fino a poco tempo fa l'unica rappresentante di mammifero conosciuta in questa classe di proteine [5], infatti negli ultimi dieci anni, una grande quantità di dati sperimentali è stata generata sulle proprietà strutturali e sulla sua attività catalitica [23].

La struttura cristallografica rivela un enzima omodimerico il cui centro del monomero è caratterizzato da un β foglietto a spirale che comprende 11 mixed strands, ed è circondato da 24 α eliche di varia lunghezza. Le due eliche $\alpha 18$ e $\alpha 19$ presentano vari residui idrofobici che costituiscono il dominio di attacco della FAAH alla membrana (Fig.4) [23]. Esperimenti di transfezione del gene della FAAH in cellule Cos7, hanno dimostrato come essa sia fortemente associata con il lato citosolico delle membrane del reticolo endoplasmatico e dell'apparato del Golgi [24].

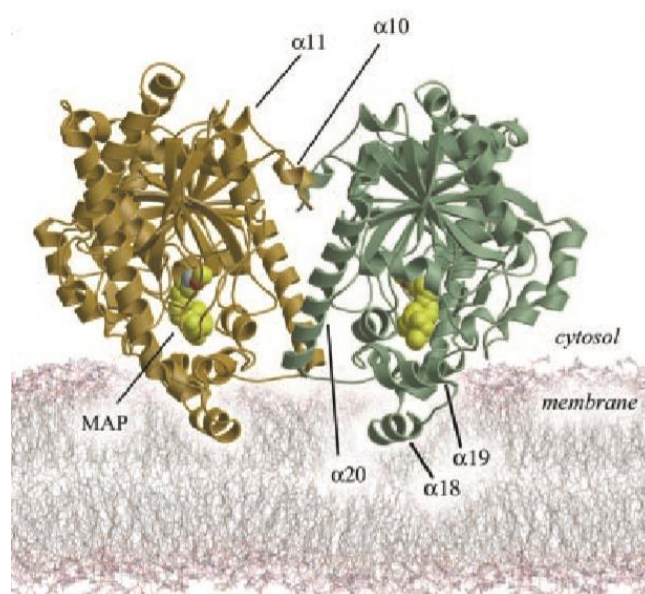


Fig.4 Struttura del monomero della FAAH [25]

Il sito catalitico della FAAH localizzato nella cavità centrale è definito dalla triade catalitica Ser241-Ser217-Lys142 ed è caratterizzato da tre canali vincolanti (Fig.5):

“Membrane access channel” (MAC), un’apertura sulla superficie della proteina che si rivolge verso il doppio strato lipidico. È adiacente alle eliche $\alpha 18$ e $\alpha 19$ e consente al substrato di entrare direttamente nel sito attivo. Vi si trovano anche i due residui carichi (Asp403 e Arg486) che possono favorire l'ingresso dei gruppi polari [23] [25];

“Acyl-chain binding pocket” (ABP), interagisce con la catena acilica del substrato durante la reazione catalitica. Nella porzione prossimale del sito attivo, il MAC e l’ABP appaiono amalgamati in un’unica indefinita sezione ed insieme formano un largo canale dalla natura prevalentemente idrofobica da una parte e moderatamente polare dall’altra [25], dove vi si trova una piccola cavità chiamata Oxianion Hole, formata dai residui Ile238, Gly239, Gly240 e Ser241, che mantiene il substrato correttamente orientato per idrolisi grazie alla sua capacità di stabilizzare la carica negativa dell’atomo di ossigeno dell’ammide durante la formazione dell’intermedio tetraedrico;

“Cytosolic Port” (CP), si trova strettamente collegata alla regione catalitica. Un secondo considerevole tunnel che emerge dal sito attivo a circa 80° dalla cavità riempita dal substrato. Questo tunnel biforca per creare una porta citosolica esposta al solvente e un

percorso bloccato dal Trp 445, un residuo che forma contatto con la membrana. Pertanto, il sito attivo della FAAH sembra accedere contemporaneamente sia all'ambiente acquoso del citoplasma che all'ambiente del doppio strato lipidico. Il citosolic port può fornire una via di uscita per i sostituenti amminici polari liberati dai substrati e potrebbe anche fornire l'ingresso per una molecola di acqua necessaria per deacilazione dell'intermedio acil-FAAH [26];

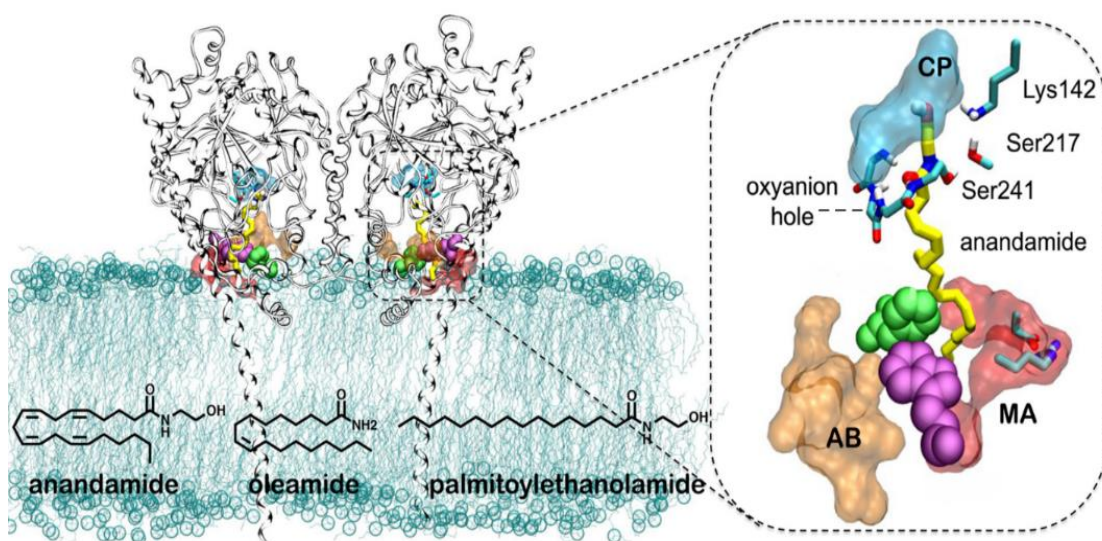


Fig.5 Sito catalitico della FAAH [26]

1.2.2 Meccanismo d'azione della FAAH

Il meccanismo catalitico della FAAH è stato studiato attraverso analisi di mutagenesi, sottolineando il ruolo della Ser241 come nucleofilo. È interessante notare come la mutazione di ogni residuo conservato di Hys non abbia portato a enzimi inattivi, indicando come la FAAH non espliciti la sua attività attraverso la triade Ser-His-Asp, tipica delle altre serine-idrolasi. Come detto più volte, la catalisi coinvolge infatti, una atipica triade Ser-Ser-Lys (Ser241-Ser217-Lys142) [15].

Le prime ricerche su un probabile meccanismo della FAAH videro il ruolo della Lys142 come catalizzatore acido/base generale con il compito di attivare il nucleofilo ed il gruppo uscente amminico [26] [27]. Sebbene il reale contributo della Ser217 invece non era ancora

chiaro, ulteriori studi dimostrarono svolgere un ruolo chiave nell'azione idrolitica della FAAH, dato che la mutagenesi FAAH S217A portò una diminuzione dell'attività enzimatica di due o tre ordini di grandezza. Infine sulla base dell'organizzazione strutturale della triade catalitica, nonché sulla misura dell'attività, della forza nucleofila, della dipendenza dal pH e della selettività amidasica/esterasica di tre FAAH mutate, fu proposto un meccanismo d'azione [5]:

- 1) *La Ser217 e la Lys142 cooperano per attivare il residuo nucleofilo Ser241. In questo scenario, la Ser217 funge da ponte tra la Lys142 e la Ser241. [28].*
- 2) *Il meccanismo coinvolge la formazione di un intermedio tetraedrico, formato dall'attacco nucleofilo della Ser241 catalitica sul gruppo carbonilico del substrato.*
- 3) *Il trasferimento concertato successivo di un H^+ tra la Lys142 e l'N-atomo del substrato permette il rilascio del gruppo amminico con il conseguente collasso dell'intermedio tetraedrico. La Ser217 funge, ancora una volta, come ponte tra i residui.*
- 4) *Infine la reazione termina con una deacilazione dell'enzima, mediata dall'ingresso di una molecola d'acqua attraverso il Citosolic Port, con il ripristino dell'attività enzimatica ed il rilascio dell'acido grasso libero.*

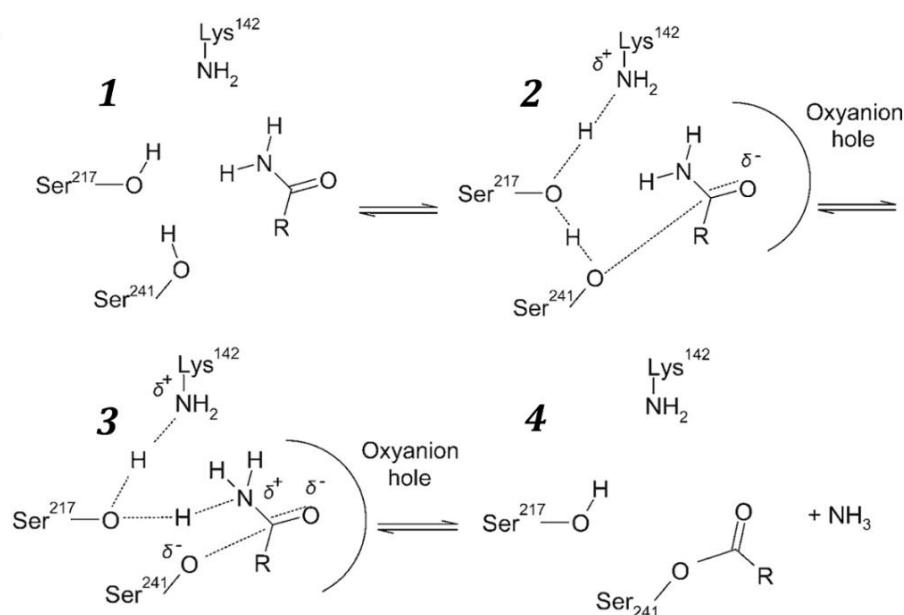


Fig.6 Meccanismo d'azione della FAAH [3]

Il primo step risulta fondamentale poiché sembra costituire la chiave teorica di come la FAAH abbia acquisito la capacità di idrolizzare e avere accesso ai substrati ammidici, meno reattivi, con la stessa efficienza e velocità dei più reattivi e abbondanti composti esterei strutturalmente correlati.

1.3 Il Sistema Endocannabinoide Come Target Terapeutico

Il sistema endocannabinoide è stato visto modulare una vasta gamma di processi fisiologici del sistema nervoso centrale e periferico, tra cui dolore, infiammazione, disturbi dell'umore e d'ansia, disturbi neurodegenerativi, cancro, aterosclerosi, infarto del miocardio, epilessia, ictus, ipertensione, glaucoma, l'obesità e sindrome metabolica, e l'osteoporosi. Nonostante l'ampio range farmacologico dei ligandi endogeni, i loro effetti collaterali psicomimetici e i potenziali abusi, hanno smorzato l'entusiasmo per lo sviluppo di farmaci che agiscono direttamente sui recettori cannabinoidi, passando quindi ad una soluzione meno invasiva attraverso studi sul loro metabolismo.

1.3.1 Trattamento del dolore e dell'infiammazione

Uno dei primi usi della cannabis è stato il trattamento del dolore. Documenti storici rivelano l'uso della cannabis in anestesia chirurgica nell'antica Cina e per alleviare il dolore di diversa natura nell'antico Israele, Grecia, Roma, e India [37]. Anandamide, THC, cannabidiolo, e cannabinoidi sintetici come CP55,940 e Win 55,212-2 sono efficaci contro alterazioni dell'organismo da parte di stimoli chimici [38], meccanici [39], e termici [40]. Infatti recenti studi su animali indicano come sia l'Anandamide che i ligandi cannabinoidi siano molto efficaci anche contro il dolore cronico sia neuropatico [41] che di origine infiammatoria [42].

Inoltre, endocannabinoidi e cannabinoidi sintetici esercitano effetti antinocicettivi sinergici in combinazione con farmaci anti-infiammatori non steroidei di uso comune, che possono avere utilità nella farmacoterapia del dolore [43], attraverso meccanismi complessi che coinvolgono effetti sul sistema nervoso centrale [44], midollo spinale [45], e nervi sensoriali periferici [46]. Ciò è coerente con la posizione anatomica dei recettori CB1 localizzati nel

distretto pertinente al dolore del cervello, nei gangli spinali e nei neuroni afferenti periferici [47]. Oltre al ruolo dei recettori CB1, ci sono prove recenti dell'implicazione dei recettori CB2 nell'attività antiperalgesica dei cannabinoidi, in modelli di dolore neuropatico acuto e cronico, soprattutto di origine infiammatoria [48]. Numerosi altri studi sull'uomo sono in corso per determinare l'efficacia del THC o estratti a base di cannabis contro le varie forme di dolore [49].

Molteplici studi sostengono l'importante ruolo del sistema endocannabinoide nella modulazione della funzione immunitaria e nell'infiammazione [50]. In primo luogo, i recettori dei cannabinoidi sono presenti sulle cellule immunitarie, dove la loro espressione è modulata da antigeni microbici o da altri stimoli che inducono l'attivazione immunitaria. In secondo luogo, la stimolazione di cellule immunitarie da tossine batteriche, come lipopolisaccaridi (LPS), aumenta i livelli cellulari degli endocannabinoidi e dei loro enzimi degradanti. In terzo luogo, gli agonisti dei cannabinoidi modulano la funzione immunitaria sia in vitro che in vivo attraverso meccanismi recettore-dipendenti e recettore-indipendenti. Gli effetti anti-infiammatori dei cannabinoidi sono complessi e possono includere la modulazione di citochine (per esempio, TNF-alfa, IL-12, IL-1, IL-6 e IL-10) e chemochine (ad esempio, CCL2, CCL5, CXCL8, e CXCL10), la modulazione del segnale adenosinico (Carrier et al., 2006), l'espressione di molecole di adesione (per esempio, ICAM-1, molecola di adesione-1 P-intercellulare e P-selectina), e la migrazione, la proliferazione e l'apoptosi di cellule infiammatorie [51].

1.3.2 Trattamento dei disordini del Sistema Nervoso Centrale

Il ruolo emergente del sistema endocannabinoide in una varietà di disturbi del sistema nervoso centrale non dovrebbe essere una sorpresa, dato l'elevato livello di espressione dei recettori CB1 nel cervello. La densità particolarmente elevata dei recettori CB1 nella corteccia, cervelletto, ippocampo e gangli della base hanno portato l'attenzione dei ricercatori alle malattie che colpiscono disturbi del movimento, dell'umore e ansia, e le condizioni relative ai meccanismi alterati di ricompensa cerebrale, così come i processi della memoria e dell'apprendimento. Infatti si è visto che il sistema endocannabinoide svolge un ruolo importante nella protezione nel caso di danno neuronale acuto (per esempio, trauma

cranico, ictus, e l'epilessia) ed anche in malattie neurodegenerative croniche, come la sclerosi multipla, il morbo di Parkinson, la malattia di Huntington e il morbo di Alzheimer [52]. Sebbene i meccanismi alla base non siano pienamente compresi, molteplici studi recettore-dipendente e recettore-indipendente sono stati seguiti per una miglior conoscenza farmacologica. Tra questi troviamo:

- La modulazione delle trasmissioni glutamatergiche eccitatori e la plasticità sinaptica tramite i recettori CB1 presinaptici [53];
- La modulazione delle risposte immunitarie ed il rilascio di mediatori infiammatori attraverso i recettori CB1, CB2 su neuroni, astrociti, microglia, macrofagi, neutrofili e linfociti [54];
- L'attivazione delle vie di segnalazione citoprotettive [55], come la proteina chinasi B/Akt [56], la proteina chinasi A [57], o fattori neurotrofici [58];
- La modulazione dell'eccitabilità e dell'omeostasi attraverso effetti sui canali del Ca², K e Na, N-metil D-aspartato (NMDA), giunzioni e intracellulare [59];
- Le proprietà antiossidanti dei cannabinoidi [60];
- L'ipotermia mediata dai recettori CB1, attraverso la riduzione del tasso metabolico e della domanda di ossigeno [61].

1.3.2.1 Trauma cranico

Ad oggi il trauma cranico è una delle principali cause di disabilità e mortalità in individui giovani, ma la terapia disponibile purtroppo è molto limitata [62]. Il trauma cranico è caratterizzato da edema cerebrale, danno neuronale, aumento della permeabilità della barriera emato-encefalica e da modifiche post-traumatiche in funzioni cognitive e neurologiche [63]. L'effetto neuroprotettivo dei cannabinoidi è stato indicato in studi effettuati in un modello di ratto affetto da trauma cranico, con il cannabinoide non psicotropo dexamabinol (HU-211). Questi studi hanno dimostrato un danno cerebrale ridotto ed un miglioramento nelle funzioni motorie e cognitive [64].

1.3.2.2 Sclerosi Multipla

La sclerosi multipla (SM) è una complessa malattia infiammatoria immuno-mediata, della sostanza bianca del cervello, che compromette la conduzione dell'impulso a causa della perdita della guaina mielinica dei neuroni [65]. L'applicazione terapeutica della cannabis nella SM è un argomento di recente dibattito [66]. Lyman et al. (1989) hanno esaminato gli effetti del THC per via parenterale in un ratto con encefalomyelitis autoimmune (EAE), un modello di laboratorio della SM. Il trattamento con THC non solo ha ridotto l'infiammazione del sistema nervoso centrale, ma ha anche migliorato la sopravvivenza rispetto al placebo.

1.3.2.3 Morbo di Parkinson

Il coinvolgimento degli endocannabinoidi nella regolazione centrale delle funzioni motorie e dei disturbi del movimento si basa su più fronti:

- I recettori CB1 sono altamente espressi nei gangli della base, soprattutto nelle aree coinvolte nel controllo motorio, substantia nigra e cervelletto [67]. Oltre ad essere anche abbondanti in queste regioni del cervello [68];
- I cannabinoidi endogeni, di origine vegetale, e sintetici hanno effetti potenti, per lo più inibitorio, sulla attività motoria [69];
- I livelli dei recettori cannabinoidi CB1 e dei loro ligandi endogeni sono alterati nei gangli della base, sia in modelli sperimentali [70] che in quelli umani, soggetti a disturbi del movimento [71];
- Il sistema endocannabinoide interagisce con diverse vie metaboliche dei neurotrasmettitori a vari livelli nel circuito dei gangli della base [72].

Il morbo di Parkinson (PID) è la seconda malattia neurodegenerativa più comune di insorgenza in età adulta, con un'incidenza di 19/100,000 persone in tutto il mondo [73]. PID è causato da una grave perdita di neuroni dopaminergici nella substantia nigra pars reticulata (SNr), con conseguente riduzione dei livelli di dopamina e una perdita di neurotrasmissione dopaminergica nello striato, che interferisce con la funzione motoria e sulla coordinazione. Esiste un ampio consenso sul fatto che il sistema endocannabinoide diventi iperattivo a livello dei gangli della base nei pazienti aventi il Parkinson [74]. Se l'aumentato livello del recettore CB1 nello striato è visto come un tentativo del cervello carente di dopamina a

normalizzare la sua funzione, l'amplificazione farmacologica di questo fenomeno fisiologico potrebbe alleviare i sintomi del morbo di Parkinson, per esempio, riducendo il rilascio di glutammato striatale [75]. Il trattamento con agonisti dei recettori CB1 può ridurre il tremore associato con iperattività del nucleo subtalamico [76] e può anche migliorare il deficit motorio [77], oltre ad inibire l'apoptosi delle cellule dopaminergiche [78]. Tuttavia, l'uso di agonisti dei recettori cannabinoidi è ancora sotto studio e risulta improbabile per ridurre la bradicinesia nel PID a causa del loro profilo ipocinetico, sia nei primati che nell'uomo [79].

1.3.2.4 Epilessia

Se l'equilibrio dei neurotrasmettitori inibitori/eccitatori viene perturbato, l'intensità della trasmissione eccitatoria può superare una certa soglia, che porta a crisi epilettiche. La stimolazione dei neuroni postsinaptici è nota per innescare la sintesi su richiesta degli endocannabinoidi attraverso un aumento di calcio e/o la stimolazione dei recettori metabotropici intracellulari [80]. Successivamente, gli endocannabinoidi vengono rilasciati e raggiungono i recettori CB1 presinaptici i quali sono in grado di portare sia un effetto inibitorio GABAergico e che eccitatorio glutamatergico tramite molteplici meccanismi [81]. La cannabis è stata usata per trattare l'epilessia per diversi secoli. I benefici della cannabis nell'epilessia sono stati riportati anche da un neurologo britannico (Reynolds, 1890), ma l'uso medico della cannabis è stata vietata nel 20esimo secolo in molti paesi. Come accaduto per diverse patologie, dopo l'identificazione della struttura del THC [82], diversi gruppi hanno studiato i suoi effetti antiepilettici [83]. Sorprendentemente, in un modello di ratto indotto in uno stato epilettico attraverso l'assunzione di pilocarpina, si è visto che gli agonisti dei recettori CB1 sono stati più efficaci nel ridurre la frequenza delle crisi rispetto ad anticonvulsivanti utilizzati clinicamente, come la fenitoina o fenobarbital. Coerentemente, il blocco dei recettori CB1 aumentava la frequenza delle crisi [84].

1.3.2.5 Schizofrenia

La schizofrenia è il secondo disturbo mentale più comune con una prevalenza un tantum di circa 0,2-2% a livello mondiale [85]. Numerose teorie sono state stese per quanto riguarda l'eziologia della schizofrenia, che vanno dai processi di sviluppo o neurodegenerativi, ai fattori ambientali fino ad anomalie dei neurotrasmettitori (dopamina o glutammato) ed ai

processi infettivi o autoimmuni. Fu anche compresa l'ipotesi cannabinoide [86] secondo cui vi è un'iperattività del sistema endocannabinoide che può portare ad uno stato iperdopaminergico ed ipoglutammatergico, che può essere alla base di alcuni dei sintomi [87]. A conferma dell'ipotesi endocannabinoide uno studio in laboratorio vide che i livelli di Anandamide nel liquido cerebrospinale o nel sangue di pazienti schizofrenici era aumentato di circa il doppio [88]. Risulta abbastanza probabile quindi che il sistema endocannabinoide possa essere un nuovo bersaglio terapeutico nella schizofrenia.

1.3.2.6 Ansia e depressione

Disturbi dell'umore come ansia o disturbo di panico, depressione maggiore o disturbo bipolare (malattia maniaco-depressiva) sono molto comuni, spesso gravi e potenzialmente pericolosi. Più del 20% della popolazione adulta presenta un disturbo dell'umore a un certo punto della loro vita. Nel 15% degli individui che mostrano un disturbo depressivo maggiore, la causa di morte è il suicidio. In accordo con una previsione dell'Organizzazione Mondiale della Sanità, entro il 2020 la depressione diventerà la seconda causa di morte prematura e di disabilità in tutto il mondo [90]. Anche se sono stati fatti progressi significativi nel trattamento dei disturbi dell'umore nel corso degli ultimi decenni, il 30% della popolazione non risponde alle attuali terapie e la ricerca di nuovi approcci farmacologici continua [90]. Molti degli effetti psicologici della cannabis e THC sono bifasici e bidirezionali, a seconda della modalità di somministrazione, della dose, della personalità, del lasso di tempo, del grado di tolleranza e di vari altri fattori ambientali ed individuali [91]. Sebbene diverse ipotesi sono state proposte per spiegare gli effetti bifasici dei cannabinoidi, tra cui recettori distinti [92] o recettori CB1 neuroanatomicamente separati, queste hanno bisogno di essere confermate in studi futuri [93]. L'alto livello dei recettori CB1 nell'ippocampo, nell'amigdala e nella corteccia prefrontale, le quali sono regioni chiave nella regolazione dell'ansia, suggeriscono che il sistema endocannabinoide svolge un ruolo in questa patologia [94]. Infatti la densità dei recettori CB1 nell'ippocampo e nell'ipotalamo risulta aumentata nel trattamento antidepressivo cronico con farmaci triciclici [95] e l'amplificazione delle azioni degli endocannabinoidi da parte dell'inibitore della FAAH URB597 è stato visto produrre effetti antidepressivo-simili, senza suscitare una potenziale dipendenza [96].

1.3.2.7 Insonnia

L'insonnia è definita come la difficoltà nell'inizio, nel mantenimento, nella durata e nella qualità del sonno che provoca la compromissione fisiologica del funzionamento diurno. La causa di insonnia spesso non è nota, ma può essere una conseguenza di una malattia cronica associata a dolore o depressione. I primi studi effettuati sul sistema endocannabinoide hanno documentato che la marijuana e il THC influenzano il sonno, sia negli esseri umani [97] che in animali da esperimento [98]. Anche l'Anandamide è stata vista modulare il sonno, aumentando la fase REM in modo dipendente al recettore CB1 [99]. Interessante notare che un recente studio ha dimostrato che l'Anandamide non solo induce il sonno, ma porta anche ad un aumento dei livelli di adenosina nel prosencefalo basale. Inoltre entrambi questi effetti sono stati bloccati dall'antagonista per i recettori CB1 SR141716 [99]. Le proprietà soporifere di alcuni potenti inibitori della FAAH, come il lipide endogeno 2-ottil-bromoacetoacetato [100], potrebbero quindi essere sfruttate terapeuticamente contro l'insonnia.

1.3.2.8 Nausea e vomito

Nausea e vomito possono presentarsi come sintomi di una varietà di malattie o come conseguenze secondarie durante una chemioterapia o radioterapia. È per quest'ultima indicazione che il THC ha guadagnato l'accettazione come agente terapeutico altamente efficace, spesso anche nei casi di tolleranza ad altri farmaci più convenzionali [101]. Il coinvolgimento dei recettori CB1 è chiaramente indicato dalla capacità dell'antagonista SR141716 di invertire gli effetti del THC e degli agonisti sintetici i quali porterebbero ad un effetto antiemetico in condizioni di vomito indotto da cisplatino [102] o cloruro di litio [103]. Questi ultimi risultati suggeriscono quindi il grande coinvolgimento del tono endocannabinoide per il trattamento di nausea e vomito.

1.3.3 Trattamento di patologie cardiovascolari

Oltre alle loro funzioni neuroprotettive ed immunologiche, gli endocannabinoidi ed analoghi sintetici esercitano importanti effetti cardiovascolari. I meccanismi alla base sono complessi e coinvolgono effetti diretti nella vascolarizzazione [104] e nel miocardio [105]. Gli studi fino ad oggi indicano che i recettori CB1 sono molto più importanti dei recettori CB2 nella regolazione cardiovascolare. I recettori CB1 sono stati visti a livello del miocardio, sia nell'uomo che nel topo, dove mediano un effetto inotropo negativo [106], ed anche nei tessuti vascolari [107]. La somministrazione endovenosa di Anandamide provoca una risposta pressoria trifasica nei topi e ratti anestetizzati, distinta da [108]:

- **Fase I**, mediata dal nervo vago. Transitoria caduta della frequenza cardiaca, della contrattilità e della pressione con un aumento di tutte le resistenze periferiche;
- **Fase II**, breve risposta pressoria (fase II) associata a un aumento della contrattilità cardiaca;
- **Fase III**, prolungato effetto ipotensivo.

L'inibizione della fase I mediata da antagonisti per i recettori TRPV1 [109] e l'assenza sia della fase I che della fase II in topi TRPV1 \neg [110], implica che queste 2 fasi sono mediate dai recettori TRPV1, invece la fase III, indotta dall'Anandamide e da cannabinoidi sintetici, è mediata da recettori CB1 [111]. L'ipotensione profonda che può essere suscitata attraverso l'attivazione farmacologica dei recettori CB1 [112] ha innescato numerosi studi per indagare il ruolo del sistema endocannabinoide nell'ipotensione associata a varie forme di shock, inclusi, lo shock emorragico [113], endotossico [114], cardiogeno [115] e lo shock associato alla pancreatite necrotizzante [116]. Gli studi iniziali hanno dimostrato che l'antagonista SR141716 del recettore CB1 preveniva o invertiva l'ipotensione associata a shock emorragici, endotossici e cardiogeno [117].

L'infiammazione cronica e lo stress ossidativo sono fattori chiave nell'arteriosclerosi e nell'invecchiamento cardiovascolare e la modulazione farmacologica di questi processi potrebbe essere di beneficio terapeutico [118]. Utilizzando modelli di topo knockout per l'apolipoproteina E, Steffens et al. (2005), riferirono che la somministrazione orale di THC portava ad una notevole inibizione della progressione della malattia. Inoltre, i recettori CB2 espressi nelle cellule immunitarie erano presenti sia nelle placche arteriosclerotiche umane che in quelle di topo e le cellule linfoidi isolate da topi trattati con THC presentavano

una diminuita capacità di proliferazione e di produzione di interferone. La cosa più importante, era che tutti questi effetti erano completamente bloccati da un antagonista selettivo per il recettore CB2, suggerendo come un approccio sui recettori CB2 potrebbe essere un target terapeutico per il trattamento dell'arteriosclerosi [119].

1.3.4 Trattamento dell'asma

L'effetto della marijuana sulle funzioni delle vie aeree è stato tra i primi ad essere esplorato come potenziale beneficio terapeutico [120]. È stato visto come fumare marijuana e l'ingestione di THC aumenti la dilatazione delle vie aeree nei soggetti sani [121] e questi effetti durino più a lungo rispetto all'effetto dei classici broncodilatatori adrenergici, come l'isoproterenolo. La broncodilatazione indotta da marijuana e dall'assunzione orale di THC è stata documentata anche in soggetti con asma lieve/moderata e nei pazienti asmatici con broncocostrizione indotta da esercizio fisico [121]. Più recentemente, Calignano et al. (2000) hanno riportato che i recettori CB1 sono presenti sui terminali degli assoni che innervano la muscolatura liscia bronchiale ed hanno visto che l'Anandamide inibisce il broncospasmo capsaicina-indotto. In conclusione, gli effetti dei cannabinoidi sulla funzione respiratoria sono piuttosto complessi e le prove per il loro potenziale terapeutico in patologie asmatiche è inequivocabile.

1.3.5 Trattamento del cancro

Gli effetti palliativi dei cannabinoidi nei pazienti oncologici sono ben noti e possono includere la stimolazione dell'appetito, l'inibizione di nausea e vomito associati a chemioterapia o radioterapia, il sollievo dal dolore, d'umore e da insonnia [122]. il Δ^9 -THC (dronabinol, Marinol) ed il suo derivato sintetico, nabilone, sono stati approvati dalla Food and Drug Administration per controllare la nausea nei pazienti oncologici sottoposti a chemioterapia e per stimolare l'appetito nei pazienti con AIDS [123]. Numerosi studi recenti hanno suggerito che i cannabinoidi possono inibire direttamente la crescita del cancro [124]. I meccanismi proposti sono complessi e possono comportare l'induzione ad apoptosi nelle

cellule tumorali, l'azione antiproliferativa ed un effetto antimetastatico attraverso l'inibizione dell'angiogenesi e della migrazione delle cellule tumorali [125]. Diversi cannabinoidi, tra cui il cannabidiolo, Anandamide e 2-AG e inibitori del trasporto degli endocannabinoidi, hanno dimostrato indurre la morte cellulare per apoptosi e di inibire la proliferazione e la migrazione in numerose linee cellulari tumorali murine e umane tra cui il glioma (C6, U87, U373, e H4), l'oligodendroglioma (GOS3), il glioblastoma multiforme, l'astrocitoma (U373-MG, U87MG), il neuroblastoma (N18TG2 e CHP100), il feocromocitoma (PC12), il tumore al seno (MCF-7, EFM-19, T47D, TSA-E1, e MDA-MB-231), il cancro della prostata (LNCaP, DU145 e PC3), il carcinoma del colon (SW 480), il carcinoma della cervice uterina (CxCa), il cancro alla tiroide (KiMol), la leucemia (CEM, HEL-92 , linee HL60, e Jurkat cellulari) e i tumori linfoidei (EL-4 e P815) [126].

In netto contrasto con quanto scritto sopra, Hart et al. (2004) hanno dimostrato che il trattamento del cancro al polmone (NCI H292), del carcinoma della pelle a cellule squamose (SCC-9), del carcinoma della vescica (5637), del glioblastoma (U373-MG), dell'astrocitoma (1321N1) e del cancro del rene (A498) con concentrazioni nanomolari di cannabinoidi, come il THC, l'Anandamide, HU-210 e il Win 55,212-2 porta ad un rapido aumento della proliferazione delle cellule tumorali. Tuttavia, lo stesso studio ha anche documentato che a concentrazioni micromolari gli stessi cannabinoidi inducono l'apoptosi delle cellule tumorali, in accordo con le precedenti relazioni. Questi risultati evidenziano un'azione bimodale dei cannabinoidi sulla crescita delle cellule tumorali.

1.4 Inibitori della FAAH presenti in letteratura

Come appena discusso, il potenziale terapeutico del sistema endocannabinoide è di notevole interesse anche se ancora non completamente esplorato. Dai vari studi sopra elencati, il minimo comun denominatore che si può ricavare, è che un'azione diretta sul tono endocannabinoide, che sia agonista o antagonista, porta ad effetti collaterali marcati, quindi l'utilizzo di inibitori selettivi verso gli enzimi responsabili della degradazione dei ligandi endogeni, Anandamide e 2-AG, è emerso come un approccio favorevole, in grado di aumentare il tono endocannabinoide solo nei tessuti dove un tale aumento sia richiesto e allo stesso tempo di ridurre gli effetti collaterali indesiderati.

Fino ad oggi numerose serie di inibitori sono state riportate in letteratura, basate soprattutto su studi estensivi di relazione struttura-attività rivolti a migliorare la potenza inibitoria. Essi possono essere suddivisi in due classi principali:

- Inibitori irreversibili
- Inibitori reversibili, che a loro volta si suddividono in:
 - *Inibitori che si legano formando un legame covalente reversibile con il sito attivo*
 - *Inibitori non covalenti*

I primi studi che hanno seguito la caratterizzazione iniziale della FAAH hanno condotto al design di inibitori competitivi quali aldeidi, α -chetoammidi, α -chetoesteri, e trifluorometilchetoni strutturalmente correlati ai substrati endogeni dell'enzima. Questi composti possiedono un carbonio elettrofilo capace di legarsi covalentemente alla Ser241 catalitica, formando un legame emiacetalico, stabilizzato dall'Oxianion Hole, che in seguito andrà incontro ad idrolisi spontanea. La sostituzione del motivo trifluorometilchetone con vari eterocicli ha portato allo sviluppo della più importante classe di ligandi covalenti reversibili, gli α -chetoeterocicli, dotati di buona selettività e attività nell'ordine del nanomolare. Studiati per inibire le serino proteasi [126] [127], una prima serie fu descritta dal gruppo di Dale Boger, composta da α -chetoeterocicli basati su arachidonoil- e oleoil-strutture. Attualmente nel Protein Data Bank sono presenti numerose strutture cristallografiche della FAAH complessata con inibitori α -chetoeterociclici, come il composto lead OL-135 (codice PDB 2WJ1) (Fig.7).

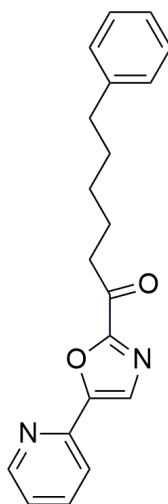


Fig.7 Struttura chimica del OL-135

Un'altra classe di inibitori ben studiata è quella dei carbammati. Questi inibitori furono studiati sempre secondo la teoria degli α -chetoeterocicli, che vedeva come requisito strutturale la presenza di un carbonile attivato, in modo da ottenere l'attacco nucleofilo da parte della Ser241 all'interno del sito catalitico. Infatti l'intermedio tetraedrico evolveva attraverso un complesso stabile acil-enzima. Questo meccanismo d'azione fu portato avanti in primo luogo da analisi di MS e da analisi ai raggi X, in studi che coinvolgevano il composto principale di questa classe di composti, l'URB597 (o KDS-4103) (Fig.8) [128]. Il composto URB597, lead della classe, presenta una porzione O-biarilica in grado di fungere da gruppo uscente nel meccanismo catalitico, con l'anello cicloesilico localizzato nell' ABP della FAAH [128]. Dopo la scoperta dell'URB597, altri analoghi furono studiati in modo da aumentare la stabilità del composto verso il metabolismo ossidativo. Infatti, dopo somministrazione in vivo si osservava un'idrossilazione della posizione C4. Tale posizione venne quindi bloccata con l'aggiunta di vari sostituenti [129], consentendo una notevole riduzione dell'idrolisi dell'Anandamide (50% a 30 nM).

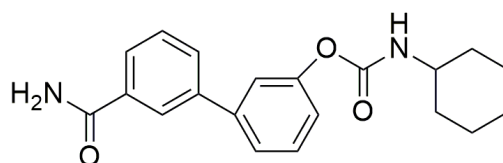


Fig.8 Struttura chimica del URB597

Oltre ai carbammati, anche derivati aril-ureidici sono stati ampiamente studiati come inibitori irreversibili della FAAH [15]. Grazie alla loro elevata resistenza all'idrolisi chimica e biologica, la funzione ureidica non sempre viene considerata come un farmacoforo. Tuttavia, è stato dimostrato che l'aggiunta di un buon gruppo uscente, ad esempio di una funzione anilinic, trasforma l'urea in una porzione più reattiva, che può quindi funzionare come inibitore dell'enzima. I derivati aril-ureidici furono scoperti da studi di HTS di librerie chimiche industriali. Sia la Janssen Pharmaceuticals che la Takeda descrissero composti sulla base di una porzione ureo-piperazinica (Fig. 9) con un valore IC₅₀ di 16 nM o 50 nM a seconda che l'enzima fosse umano o di ratto [130].

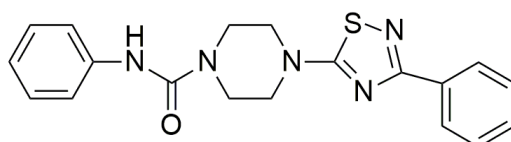


Fig.9 Struttura chimica dei derivati aril-ureidici

Recentemente, è stato segnalato un nuovo tipo di inibitori della FAAH che utilizza l'acido boronico come funzione elettrofila. Questa funzione è stata già descritta per inibire serino proteasi in modo reversibile [131]. Infatti, la capacità del boro di salire da una geometria planare trigonale ad una geometria tetraedrica permette agli acidi boronici di formare un transitorio e reversibile intermedio tetraedrico con la serina nucleofila. Sia Infinity Pharmaceuticals [132] che Minkkila e collaboratori [133] pubblicarono nel 2008, i primi acidi aril-boronici descritti come inibitori della FAAH (Fig.10). Questi due inibitori sostengono un'attività nanomolare (IC₅₀ = 9,1 nM) ed un'inibizione reversibile. Inoltre, indagini supplementari sono state intraprese per svelare le interazioni tra l'inibitore e l'enzima e sia la modellazione molecolare che studi di mutagenesi della rFAAH hanno dimostrato che il fenile di questi inibitori interagisce con il canale idrofobico dell'enzima.

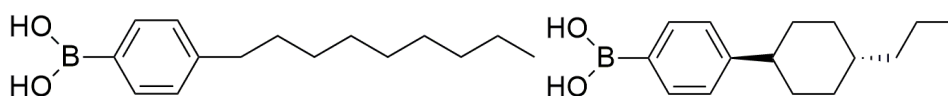


Fig.10 Struttura chimica dei derivati boronici

Nel campo abbastanza affollato degli inibitori della FAAH, molta attenzione deve essere data all'attuale selettività di questi composti. In particolare, la ridondanza del sistema endocannabinoide sia in termini di target molecolari che di enzimi metabolici suggerisce che la selettività, piuttosto che la potenza di un inibitore, deve essere un requisito obbligatorio [15]. La maggior parte degli inibitori covalenti reversibili e irreversibili mostra selettività per la FAAH nei tessuti cerebrali, ma possiede target addizionali in tessuti periferici come fegato e rene. Molti isoenzimi appartenenti alla famiglia delle carbossilesterasi, coinvolti nel metabolismo di farmaci e profarmaci, sono stati identificati come off-target [134]. Come precedentemente detto, la scoperta e lo sviluppo di inibitori prettamente non covalenti si pone come una valida strategia per cercare di migliorare il profilo di selettività e sicurezza mostrato dalle altre classi di inibitori.

Ad oggi nel Protein Data Bank esistono soltanto due cristalli della FAAH complessata con un inibitore non covalente: tali cristalli (codici PDB: 3QJ9, 3QK5) riportano i composti rappresentanti (Fig.11) della serie identificata da Gustin e collaboratori [135], nel sito attivo della proteina di ratto. L'inibizione è ottenuta primariamente per complementarità di forma con il sito attivo e attraverso numerose interazioni idrofobiche a livello dell'ABP e del MAC, che giustificano l'attività inibitoria nanomolare nei confronti della h/FAAH (IC₅₀= 10 nM per il 3QJ9, 2 nM per il 3QK5). Non è presente nessun legame a idrogeno con la proteina mentre legami a idrogeno non critici sono instaurati con alcune molecole di H₂O all'interno del sito.

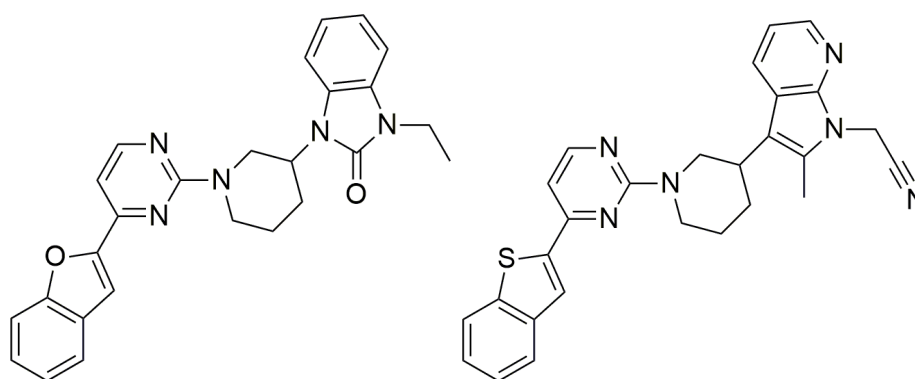


Fig.11 Struttura chimica del 3QJ9 (sinistra) e del 3QK5 (destra)

Si riportano la struttura e una breve descrizione dei composti attivi ad oggi più interessanti, in grado di inibire la FAAH attraverso interazioni non covalenti.

2_16j: appartenente ad una serie di benzotiazoli sintetizzata da Wang e collaboratori nel 2009 [136]. Grazie ai loro studi stabilirono che il legame tra la porzione benzotiazolica e l'enzima è principalmente guidato da interazioni idrofobiche, manifestate dalla minor attività degli analoghi tiazolici ma soprattutto dalla minor potenza dei derivati benzossazolici e benzoimidazolici, più polari. Il 2_16j (Fig.12) presenta un'attività ed una selettività molto elevate nei confronti della h/FAAH, con un IC₅₀ di 2 nM. Thian et al. [137] hanno dimostrato il carattere non-covalente dell'inibizione, e attraverso studi di docking hanno proposto un possibile binding mode da analogo dello stato di transizione dell'enzima: i residui catalitici Ser241 e Ser217 formano un legame ad idrogeno con un atomo di ossigeno del solfone, mentre il backbone dei residui Ile238 e Gly239 forma un legame ad idrogeno con l'altro atomo di ossigeno.

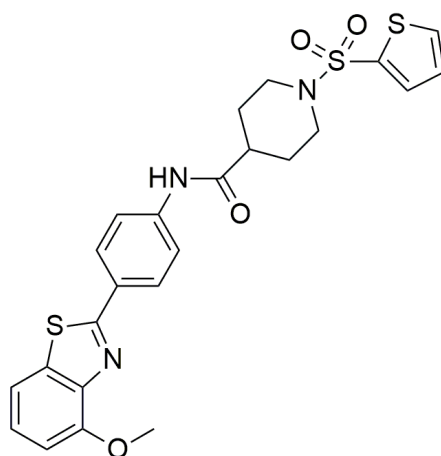


Fig.12 Struttura chimica del 2_16j

Composti 14 e 16: questi composti (Fig.13) appartengono ad una classe di tioidantoini e imidazolidindioni, sintetizzata da Muccioli e colleghi nel 2005 [138] a partire da ligandi del recettore cannabinoide CB1. Il nucleo idantoinico infatti ha molti effetti farmacologici ed è presente in diversi farmaci clinicamente rilevanti (ad esempio, nilutamida e fenitoina [139]). Opportune modifiche strutturali hanno permesso di spostare l'attività verso la FAAH, con una IC₅₀ rispettivamente di 1 μ M e 7 μ M sulla proteina di ratto. Il binding mode stabilito

attraverso studi di docking prevede la catena alchilica degli inibitori all'interno dell'ABP, mentre uno degli anelli fenilici rivolto verso la triade catalitica [140].

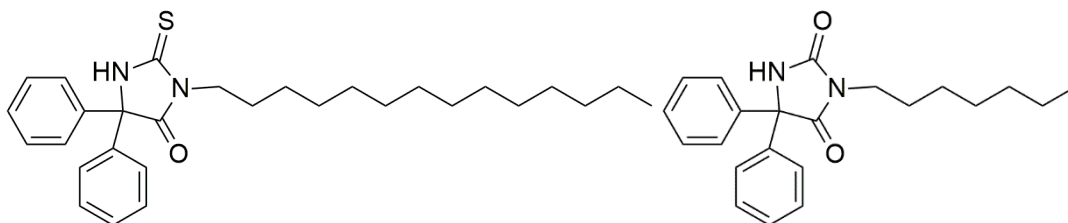


Fig.13 Struttura chimica del composto 14 (sinistra) e 16 (destra)

AZ513: questo ligando è stato identificato da uno screening sull' "Astrazeneca compound collection" utilizzando la h/FAAH e il substrato sintetico fluorescente decanoil-amminometilcumarina (D-AMC) [141].

Inibisce la FAAH umana con un $IC_{50} = 551 \text{ nM}$, 20 volte più potente rispetto all'inibizione della FAAH di ratto ($IC_{50} = 27 \text{ nM}$). È inattivo a $10 \mu\text{M}$ contro la serina idrolasi acetilcolinesterasi, la trombina, e la tripsina. Contrariamente alla maggior parte altri inibitori della FAAH, l'AZ513 non ha mostrato alcuna modifica covalente al sito catalitico dell'enzima, esibendo inoltre un'inibizione reversibile. In un saggio enzimatico trasversale con l'inibitore irreversibile OL-135, si è visto che l'AZ513 non ha partecipato competitivamente al sito catalitico della FAAH. Questo fatto indica quindi come l'AZ513 legghi in un sito differente da quello del substrato endogeno. Inoltre l'AZ513 ha una buona penetrazione cellulare come dimostra l'inibizione della FAAH in cellule HEK293 umane ($IC_{50} = 360 \text{ nm}$).

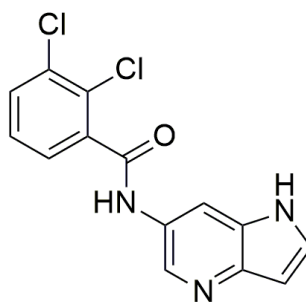


Fig.14 Struttura chimica di AZ513

JZP-327A: uno studio del 2013 di Patel e collaboratori [142] prevedeva l'identificazione di un potente, e selettivo, composto chirale che andasse ad inibire la FAAH. Il primo step fu quello di sintetizzare l'1,3,4-ossadiazol-2-one dall'ibuprofene, sostituendo poi la catena p-isobutilica con p-Ph (flurbiprofene) e con m-OPh (fenoprofene). I primi due composti mostrarono un'attività inibitoria eccellente (IC₅₀ 48nM) mentre il fenoprofene si fermava solo al micromolare. Inoltre questi composti mostravano selettività anche per l'enzima MAGL. Aperta la strada di un possibile miglioramento dell'attività attraverso analisi di SAR, i ricercatori identificarono il migliore inibitore della FAAH di ratto, umanizzata, con un IC₅₀ di appena 11nM che vedeva in para all'anello benzenico, una catena isobutilica (Fig.15). Questo JZP-327A aveva una grande selettività anche per la MAGL, le COX ed altre serino-idrolasi. Data la sua natura chirale, l'S-enantiomero risultava 20 volte più attivo del corrispondente enantiomero R. Sebbene studi meccanicistici indichino la natura non competitiva del composto, studi di docking mostrano come questo assuma interazioni favorevoli all'interno del sito attivo, incluso un importante legame ad idrogeno dell'ossigeno carbonilico con la tasca ossianionica.

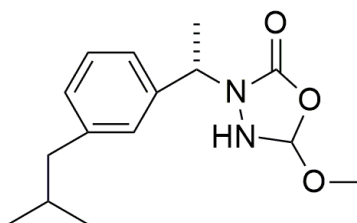


Fig.15 Struttura chimica del JPZ-327°

Composto 3_158: appartenente ad una serie di benzossazoli brevettati dalla Renovis [143], questo composto presenta un IC₅₀ di 1,2 nM sulla h/FAAH.

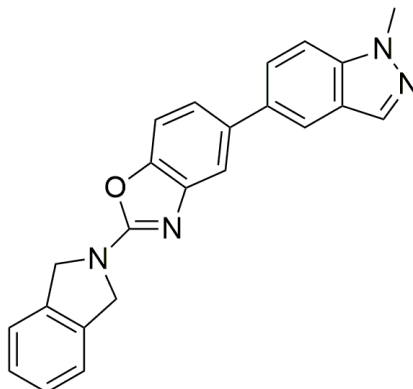


Fig.16 Struttura chimica del composto 3_158

Composto 5_68: appartenente ad una serie di ammino pirimidine/triazine recentemente brevettate dalla Janssen Pharmaceuticals, questo composto ha un IC₅₀ di 1 nM sulla h/FAAH [144].

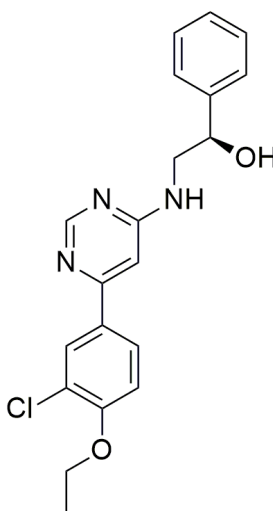


Fig.17 Struttura chimica del composto 5_68

Composto 6_101: la Merk ha sottoposto a brevetto una serie di imadazoli/ossazoli che apparirebbero agire come inibitori non covalenti della FAAH. L'IC₅₀ del composto riportato è di 6,3 nM sulla proteina umana [145].

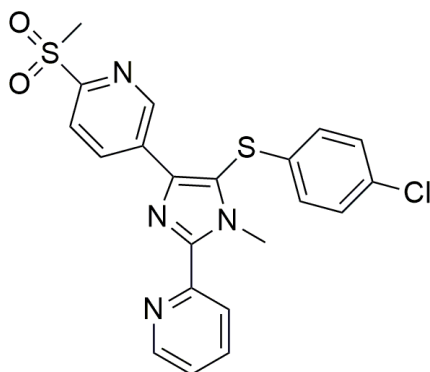


Fig.18 Struttura chimica del composto 6_10

CAPITOLO 2 METODI

2.1 Studi di Docking

La procedura di Docking è stata condotta su un inibitore non – covalente (VS8) individuato tramite studi precedenti effettuati nel nostro laboratorio. L’analisi ha visto l’utilizzo di alcuni tra i software maggiormente utilizzati nel campo della modellazione molecolare: GOLD, Glide, AutoDock, Vina, DOCK, FRED.

2.1.1 Gold

GOLD (Genetic Optimization for Ligand Docking) è un algoritmo genetico che permette il docking flessibile ed automatico di un set di composti all’interno del sito di legame di una proteina. Le molecole analizzate vengono posizionate nel sito recettoriale, individuato dal programma secondo la modalità definita dall’utente, e le varie soluzioni vengono valutate sulla base del punteggio ottenuto a seconda della scoring function utilizzata. Le migliori “n” soluzioni per ogni molecola, a seconda di quanto impostato, vengono ottenute come file di output al termine del calcolo.

Le scoring function a disposizione sono quattro: GoldScore, ChemScore, ASP (*Astex Statistical Potential*) e CHEMPLP (*Piecewise Linear Potential*). La scelta della funzione da utilizzare può essere fatta in maniera razionale, seguendo alcune linee guida.

GoldScore valuta il contributo energetico finale di una pose in base ai seguenti cinque parametri: legami a idrogeno proteina – ligando, forze di Van der Waals tra proteina e ligando, forze di Van der Waals interne al ligando, tensioni interne dovute a torsionali del ligando e legami a idrogeno intramolecolari del ligando.

ChemScore è una funzione empirica derivata dallo studio di un insieme di 82 complessi per cui era disponibile la misura di affinità di legame. ChemScore stima la variazione

totale di energia libera che si verifica durante il legame con il ligando aggiustando il valore finale con l'aggiunta di una penalizzazione in termini di torsioni interne.

ASP è paragonabile a ChemScore e GoldScore in termini di precisione. L'approccio utilizzato, nonostante ciò, è diverso: ASP raccoglie informazioni sulla frequenza di interazione tra gli atomi del ligando e quelli della proteina e confronta i risultati trovati con strutture PDB esistenti. Queste informazioni vengono utilizzate per generare potenziali statistici.

PLP, **CHEMPLP** sono funzioni empiriche ottimizzate per la previsione della pose di legame; la seconda delle due costituisce la scoring function predefinita di GOLD. In entrambi i casi, il Piecewise Linear Potential (PLP) viene utilizzato per modellare la complementarità sterica tra proteina e ligando. Entrambe le funzioni sono in grado di considerare i legami covalenti.

Per il funzionamento del programma può essere utilizzata l'interfaccia grafica, facile ed intuitiva, che consentirà di caricare la proteina (e l'eventuale ligando di riferimento) e le molecole da sottoporre alla procedura di docking, nonché di impostare i parametri da utilizzare per il calcolo. Tutte queste informazioni verranno racchiuse in un file, avente l'estensione *.conf*, il quale potrà essere modificato successivamente a seconda delle esigenze dell'operatore (come nel nostro caso).

È riportato di seguito il file "*asp.conf*", utilizzato per effettuare il docking sul ligando di riferimento VS8 (nel caso in cui è stata utilizzata la funzione *asp*). I composti sono stati "doccati" con tutte e quattro le scoring function disponibili, lasciando gli altri parametri inalterati. I più importanti sono stati accompagnati da una breve spiegazione, preceduta da cancelletto.

"GOLD CONFIGURATION FILE"

AUTOMATIC SETTINGS

autoscale = 1 # definisce il grado di efficienza nella ricerca delle soluzioni. Esso permette di controllare la velocità del docking e l'affidabilità dei risultati quando i parametri delle sezioni "population" e "genetic operators" vengono impostati su auto. Impostando il valore di 1, GOLD tenderà di applicare i settaggi ottimali per ogni ligando.

POPULATION

*popsiz*e = *auto*

select_pressure = *auto*

n_islands = *auto*

maxops = *auto*

niche_size = *auto*

GENETIC OPERATORS

pt_crosswt = *auto*

allele_mutatewt = *auto*

migratewt = *auto*

settando tutti i parametri visti in queste due categorie su *auto*, GOLD calcolerà automaticamente con il suo algoritmo genetico il numero ottimale di operazioni per ogni ligando.

FLOOD FILL

radius = 10 # il sito di legame viene definito come tutti gli atomi all'interno di 10 Å da un punto centrale specificato. Questo raggio è abbastanza grande da contenere ogni possibile binding mode del ligando all'interno del sito attivo della FAAH.

origin = 0 0 0

do_cavity = 1

floodfill_atom_no = 0

cavity_file = /home/alex/nfs/parametri/GOLD/lig.mol2 # viene riportato il path completo del ligando da utilizzare come riferimento per la definizione del sito di legame.

floodfill_center = *cavity_from_ligand* 10 *atom* # il sito di legame viene definito sulla base della posizione del ligando di riferimento. 10 *atom* istruisce GOLD a includere nel sito di legame solo gli atomi all'interno di 10 Å di distanza dal ligando, non considerando gli atomi dei residui associati fuori da tale distanza.

DATA FILES

ligand_data_file /home/alex/nfs/parametri/GOLD/lig_min.mol2 1000 # viene riportato il path completo del file dei ligandi da dockare (con idrogeni). Nel nostro caso dato che il ligando era solo uno, abbiamo prima effettuato una minimizzazione ed un'analisi conformazionale del ligando. 1000 indica il numero di cicli di algoritmo genetico a cui ogni ligando deve essere sottoposto (solitamente questo valore è 30, ma per procedure di docking più pesanti si può aumentare)

param_file = DEFAULT

set_ligand_atom_types = 1

set_protein_atom_types = 1

directory = . # cartella in cui depositare i file di output

tordist_file = DEFAULT

make_subdirs = 0

save_lone_pairs = 0

fit_points_file = fit_pts.mol2

read_fitpts = 0

FLAGS

internal_ligand_h_bonds = 0

flip_free_corners = 1 # viene consentita l'inversione conformazionale degli anelli del ligando

match_ring_templates = 0

flip_amide_bonds = 1 # viene consentita l'inversione conformazionale dei legami ammidici, tioammidici, ureidici e tioureidici

flip_planar_n = 1

flip_ring_NRR *flip_ring_NHR*

flip_pyramidal_n = 1 # viene consentita la possibilità di inversione piramidale degli atomi di azoto

rotate_carboxylic_oh = flip # viene consentita l'inversione conformazionale degli acidi carbossilici protonati

use_tordist = 1

postprocess_bonds = 1

rotatable_bond_override_file = DEFAULT

solvate_all = 1

TERMINATION

early_termination = 0

n_top_solutions = 3

rms_tolerance = 1.5

CONSTRAINTS

force_constraints = 0

COVALENT BONDING

covalent = 0

SAVE OPTIONS

save_score_in_file = 1

save_protein_torsions = 0

concatenated_output = asp.sdf

clean_up_option delete_all_solutions

clean_up_option save_top_n_solutions 5 # vengono restituite le migliori 5 soluzioni per ogni diversa conformazione del ligando

clean_up_option delete_redundant_log_files

clean_up_option delete_all_initialised_ligands clean_up_option delete_empty_directories

clean_up_option delete_rank_file

clean_up_option delete_all_log_files

output_file_format = MACCS

FITNESS FUNCTION SETTINGS

initial_virtual_pt_match_max = 6

relative_ligand_energy = 1

gold_fitfunc_path = asp # viene indicata la scoring function da utilizzare tra asp, plp, chemscore e goldscore

score_param_file = DEFAULT

PARALLEL OPTIONS

hostfile = /home/alex/nfs/parametric/GOLD/gold.hosts # describe il path dove cercare il file .hosts al cui interno c'è scritto se mandare il calcolo in più processori e nel caso, quanti processori utilizzare

PROTEIN DATA

protein_datafile = /home/alex/nfs/parametri/GOLD/prot.pdb # path completo del file contenente la proteina (con idrogeni) su cui effettuare il docking.”

Il file *asp.conf* potrà essere utilizzato per lanciare il docking direttamente da shell attraverso il comando:

- *gold_auto asp.conf*

Il comando “gold_auto” consente di lanciare il calcolo in monoprocesso. Per lanciare il calcolo in parallelo è necessario innanzitutto creare un file dove vengono indicati il nome della macchina e il numero di processori da utilizzare. Nel file *.conf* deve essere indicato il path del file creato nella sezione PARALLEL OPTIONS (vedi sopra). Nel nostro caso il file è stato nominato *gold.hosts*, ed è qui riportato:

“hyperion.farm.unipi.it 6 no_of_processes 6”

Adesso è possibile lanciare il docking in parallelo attraverso il seguente comando digitato da shell:

- `parallel_gold_auto 6 asp.conf`

dove 6 è il numero di processori da utilizzare per il calcolo (deve corrispondere al numero riportato nel file *gold.hosts*).

In questo progetto, tutti i docking eseguiti con GOLD sono stati lanciati da shell, sfruttando sia la modalità che lavora in monoprocesso che quella in parallelo.

2.1.2 *Glide*

Glide (Grid based Ligand Docking with Energetics) è un programma di docking che utilizza una serie gerarchica di filtri per la ricerca della migliore pose che il ligando può assumere nel sito attivo del recettore. La forma e le proprietà di quest'ultimo sono rappresentate su una griglia da diversi insiemi di campi che forniscono un punteggio della pose sempre più accurato.

Ogni ligando viene diviso da GLIDE in una regione centrale chiamata *core* e in un certo numero di gruppi *rotameri*, ognuno di questi legato al *core* attraverso un legame ruotabile. Durante la generazione della conformazione, ogni *core* viene rappresentato da una serie di conformazioni di base (in genere 500), il cui numero dipende dal numero di legami ruotabili e da altri parametri geometricamente limitanti. Per ogni conformazione viene effettuata una ricerca esaustiva delle possibili orientazioni all'interno del sito attivo della proteina:

Il primo passo consiste nella selezione dei cosiddetti “*site points*”, ovvero punti equidistanti 2 Å appartenenti alla griglia del sito attivo, e che verranno utilizzati per il posizionamento degli atomi.

In una seconda fase vengono scartate le conformazioni che presentano troppe repulsioni steriche con il recettore, in modo da selezionare le possibili orientazioni del diametro del ligando (inteso come la linea che collega i due atomi più lontani tra loro). A questo punto vengono valutate le interazioni di tutti gli atomi capaci di fornire legami ad idrogeno, e a queste viene assegnato un punteggio; se è buono abbastanza, allora viene dato un

punteggio a tutte le interazioni con il recettore. A questo scopo viene utilizzata la funzione Chemscore, che riconosce come favorevoli le interazioni a idrogeno e idrofobiche, mentre penalizza le repulsioni steriche.

Solo un numero ridotto di pose è in grado di procedere verso la terza fase, dove viene applicata una minimizzazione energetica sulla base di griglie elettrostatiche e di van der Waals, e del force-field OPLS-AA. Infine, le pose minimizzate sono rivalutate utilizzando *GlideScore*, una funzione di scoring che si basa su ChemScore ma che differisce per alcuni termini, come ad esempio il calcolo degli squilibri elettrostatici. La scelta della pose migliore per ogni ligando è data dalla funzione *Emodel*, che combina il punteggio della griglia energetica, quello di *GlideScore*, e l'energia interna della molecola.

Quella appena vista era la funzione standard di GLIDE, denominata appunto SP (Standard Precision), ma è a disposizione anche un'altra funzione più avanzata ed accurata chiamata XP (Extra Precision). Essa è stata progettata per identificare le pose potenzialmente sfavorevoli dal punto di vista chimico fisico, in modo da ridurre il numero di falsi positivi e fornire una buona correlazione tra le pose migliori e i punteggi migliori.

Per poter effettuare il docking con questo programma, è necessario innanzitutto definire la griglia recettoriale che sarà utilizzata come riferimento per l'assegnazione del punteggio alle soluzioni analizzate. Per creare la griglia è sufficiente accedere alle funzioni di GLIDE dall'interfaccia grafica di MAESTRO (essendo un programma della suite Schrodinger), caricarvi il file del complesso proteina-ligando di riferimento (in questo caso la proteina 3OJ8 con il ligando VS8), eliminare l'acqua eventualmente presente ed attivare la funzione "*Receptor grid generation*". In questo caso la griglia relativa al sito catalitico della FAAH è stata definita sulla base della posizione del ligando VS8 complessato, in un intorno di 10 Å dal punto corrispondente al centro del ligando stesso.

Adesso è necessario accedere alla funzione "Ligand docking", dove verrà caricato il file della griglia (grid.zip) e sarà scelta la funzione di docking desiderata (XP o SP; in questo progetto sono state utilizzate entrambe). Nella sezione "Ligands" verrà indicato il file di molecole da sottoporre alla procedura, mentre nella sezione "Output" verrà scelta l'estensione del formato in uscita e il numero di pose da salvare per ogni ligando. Nel nostro caso si è deciso di ritenere soltanto 1 pose (come di default), che sarà inoltre minimizzata dopo il calcolo di docking ("Perform post docking minimization").

2.1.3 AutoDock

AutoDock è un programma di docking che utilizza un algoritmo genetico, il *Lamarckian Genetic Algorithm*, per il calcolo della pose migliore che interagisce con il sito attivo della proteina. Dopo aver calcolato inizialmente una popolazione di possibili soluzioni, l'algoritmo ne selezionerà una parte in base alle funzioni di scoring e darà origine a una nuova popolazione di soluzioni figlie, da cui avrà inizio un secondo ciclo di generazione e così via. In questo modo il “genotipo”, ovvero la stringa binaria a cui corrisponde ciascun ligando, verrà influenzato da fattori esterni, esattamente come nell'ipotesi lamarckiana.

Le popolazioni di soluzioni sono ottenute tramite operatori genetici (mutazioni, crossover e migrazioni) che imitano quelli biologici. I gradi di libertà sono codificati in geni o stringhe binarie, e a geni e cromosomi è assegnato un valore basato sulla fitness della scoring function. Le operazioni di mutazione causano cambiamenti nel valore di un gene, mentre il crossover muove un set di geni da un cromosoma “genitore” ad un altro; la migrazione invece muove singoli geni da una sottopopolazione ad un'altra.

L'interazione tra ligando e recettore è valutata in due fasi, calcolando la variazione di energia intramolecolare del passaggio dalla forma libera a quella legata e la variazione di energia libera intermolecolare implicata nello stesso passaggio.

2.3.3.1 Preparazione dei file di input

Per effettuare il docking è necessario per prima cosa preparare le coordinate di ligando e recettore. La preparazione delle coordinate è la fase più importante nella procedura, poiché in esse sono inclusi parametri fondamentali come: idrogeni polari, atom – type e cariche parziali. Le coordinate del ligando originale e della macromolecola sono trattate separatamente ed i loro file sono in un formato particolare, il PDBQT.

Per creare le coordinate abbiamo utilizzato l'interfaccia grafica di AutoDock, chiamata *AutodockTools*. Usando l'interfaccia grafica è possibile creare le coordinate in modi diversi: il più semplice e veloce è sicuramente quello di andare sulla casella “Grid” ed aprire il file (del ligando o della proteina). AutoDock assegnerà automaticamente le cariche parziali, idrogeni ed atom type scrivendo tutto in un file PDBQT. Il programma ci comunicherà inoltre il numero di idrogeni non polari da fondere con i rispettivi carboni.

2.3.3.2 Generazione della griglia

Per generare la griglia necessaria ad Autodock come riferimento per effettuare il docking è necessario calcolare delle mappe di parametri. Ciò viene fatto grazie al programma AutoGrid4, al quale devono essere forniti i parametri da utilizzare per la creazione della griglia, ovvero il recettore che deve essere mappato, il tipo di mappe, la dimensione e la posizione di tali mappe. In generale è calcolata una mappa recettoriale per ogni atom-type del ligando che deve essere doccato, una per il potenziale elettrostatico e una per la desolvatazione.

Per la definizione dei parametri necessari per la creazione della griglia abbiamo utilizzato l'interfaccia grafica, procedendo nel seguente modo:

- Importare i file PDBQT generati in precedenza della proteina e del ligando complessato;
- Dalla barra dei menu scegliere *Grid/Set Map Types/Directly*. Scegliendo questa opzione si apre il widget *AutoGpf Ligand* che rende possibile la modifica delle mappe da calcolare. Cliccare su *Accept* per chiudere il widget;
- Selezionando dal menu *Grid/Grid Box* si apre la finestra *Grid Option*. Questa possiede dei bottoni in alto: *File*, *Center*, *View* e *Help*. Il tasto *File* consente di chiudere la finestra salvando o meno le modifiche; il tasto *Center* serve per centrare il box della griglia; il tasto *View* è utile per modificare la visibilità del box. Nella finestra di dialogo sono inoltre mostrati i *Current Total Grid Points* delle mappe che definiscono le dimensioni di ognuna. È possibile agire variando le dimensioni delle mappe, le distanze tra i punti della griglia e la posizione del centro di essa. Una volta modificata la griglia come desiderato cliccare su *File/Close saving current*;
- Selezionare dal menu *Grid/Output/Save GPF* per salvare un file con i parametri della griglia. Si utilizza l'estensione GPF.

Per la creazione della griglia abbiamo seguito la seguente procedura:

- Cliccare su *Run/Run Autogrid*; nella stessa finestra specificare il computer su cui si desidera lavorare cliccando *Run/Host Preferences*. Specificare il percorso cartella dei file eseguibili e dei file contenenti i parametri, e il nome che verrà dato ai file di output nell'apposito spazio;
- Una volta impostate tutte le preferenze cliccare dal menu il tasto *Launch* per lanciare il calcolo.

2.3.3.3 Inizio del calcolo di docking

Per poter iniziare il docking è necessario trasformare ulteriormente i ligandi dal formato PDBQT al formato DPF, che non contiene solo le coordinate ed i parametri dei ligandi ma anche altre informazioni importanti come il numero di step che devono essere effettuati dall'algoritmo durante il docking. Abbiamo effettuato questa operazione in maniera automatizzata utilizzando lo script *adt_dock*. La stringa che permette la conversione di un singolo file è:

- `prepare_dpf4.py -l input_lig.pdbqt -r input_prot.pdbqt -p ga_run=200 -o output.dpf`

Dove `-l`= file di input del ligando; `-r`= file della di input della proteina; `-p`= seguito dal numero di step effettuati dall'algoritmo genetico; `-o`= indica il nome del file di output.

È importante scegliere adeguatamente il numero di step da far effettuare all'algoritmo in quanto non solo influenza l'accuratezza del calcolo ma anche il tempo macchina.

Avendo preparato tutti i file necessari si può procedere con il docking, per farlo si digita nella shell di comando:

- `Autodock4 -p output.dpf -l output.glg`

Dove `-p`= file di input contenente i parametri necessari all'algoritmo per effettuare il docking del ligando; `-l`= file di input contenente i parametri della griglia.

I risultati del docking sono in formato DLG, ed è necessario convertirli in un formato adatto per il processamento. Per farlo è stato utilizzato il programma *write_lowest_energy_ligand.py*, che è in grado di generare un file PDBQT contenete la pose

del ligando restituita come migliore. È stato creato uno script, chiamato *write_BE*, in grado di automatizzare la procedura per tutti i ligandi. A questo punto è possibile procedere direttamente al secondo passaggio del processamento, in quanto i file sono già nel formato corretto.

2.3.3.4 Autodock da shell

Autodock, come gli altri programmi di docking da noi utilizzati, oltre ad un uso da interfaccia grafica, preveder anche un uso da shell, più diretto, veloce e comodo. Gli step da fare sono sempre i soliti:

Preparazione file di input del ligando attraverso l'input:

- *Pythonsh/usr/local/MGLTools-1.5.6rc3/MGLToolsPckgs/AutoDockTools/Utilities24/prepare_ligand4.py -l lig.mol2 >> log*

Preparazione file di input della proteina attraverso l'input:

- *Prepare_receptor4.py -r prot.pdb -o prot.pdbqt*

Generazione delle coordinate del box e degli atom type da inserire nella griglia:

```
#!/bin/csh -f
```

```
./pdbbox_15 lig.pdbqt > log
```

```
pythonsh/usr/local/MGLTools-
```

```
1.5.6rc3/MGLToolsPckgs/AutoDockTools/Utilities24/prepare_gpf4.py -l lig.pdbqt -r  
prot.pdbqt -p ligand_types=A -vcap prot.gpf | sed 's/ligand_types A/ligand_types A C F  
HD N NA OA S P SA Br Cl I/' > TEMP
```

```
cat TEMP | sed '10i\map prot.I.map\'/sed '10i\map prot.Cl.map\'/sed '10i\map  
prot.Br.map\'/sed '10i\map prot.SA.map\'/sed '10i\map prot.P.map\'/sed '10i\map  
prot.S.map\'/sed '10i\map prot.OA.map\'/sed '10i\map prot.NA.map\'/sed '10i\map  
prot.N.map\'/sed '10i\map prot.HD.map\'/sed '10i\map prot.F.map\'/sed '10i\map  
prot.C.map\' > prot.gpf
```

```
setenv var `grep "gridcenter" log`
```

```
set gct=`echo $var|sed 's/ #xyz-coordinates or "auto"/"/|sed 's/REMARK gridcenter //'`
```

```
set c1=`echo $gct | awk '{print $1}'`
```

```
set c2=`echo $gct | awk '{print $2}'`
```

```
set c3=`echo $gct | awk '{print $3}'`
```

```
cat prot.gpf | sed 's/gridcenter auto/gridcenter '$c1' '$c2' '$c3'/' > pro
```

```
cat pro | sed 's/././g' > prot.gpf
```

```
setenv var1 `grep "npts" log`
```

```
set gct1=`echo $var1|sed 's/ # num.grid points in xyz"/|sed 's/REMARK npts //'`
```

```
set d1=`echo $gct1 | awk '{print $1}'`
```

```
set d2=`echo $gct1 | awk '{print $2}'`
```

```
set d3=`echo $gct1 | awk '{print $3}'`
```

```
sed 1d prot.gpf > int
```

```
sed '1i\npts '$d1' '$d2' '$d3'\` int > pro1
```

```
mv pro1 prot.gpf
```

```
rm int
```

```
rm pro
```

```
rm log
```

./pdbbox 15 = assicurarsi di avere il programma PDBBOX15 all'interno della cartella dove si lancia lo script, poiché è quello che genera il box in base alla posizione del ligando. Pythonsh... = serve per creare le mappe degli atomi della proteina e trasferirle nel file di output prot.gpf, creando prima un file TEMP dal quale poi attraverso le funzioni di cat, i dati vengono inseriti nel file prot.gpf. Set = servono per scrivere le coordinate e le dimensioni della griglia nel file prot.gpf.

Una volta creato il file `prot.gpf` si può mandare il calcolo della griglia tramite il comando:

- `Autogrid4 -p prot.gpf -l prot.glg`

L'ultimo step di docking ed il suo processamento sono stati già affrontati

2.1.4 Vina

Vina rappresenta la nuova versione di AutoDock, ed infatti presenta molte similitudini con il suo predecessore ma, allo stesso tempo, anche molte differenze.

Una differenza importante consiste nella velocità di calcolo, dato che Vina è molto poco dispendioso sotto questo punto di vista; altra differenza fondamentale è rappresentata dal fatto che al momento del calcolo della griglia, Vina calcola internamente ed automaticamente le “grid maps” viste per AutoDock. Questo costituisce un grande vantaggio in termini di facilità e velocità di esecuzione. Inoltre, le funzioni di scoring e gli algoritmi utilizzati in questo tipo di analisi risultano essere completamente diversi rispetto al suo predecessore, cosa che porta a considerare Vina quasi come un software a sé stante.

Anche questo programma, che utilizza un algoritmo genetico per i calcoli di docking, necessita di file con estensione PDBQT, quindi recanti i parametri coordinata già visti per AutoDock. Essi verranno creati nello stesso modo della suite del predecessore. La griglia necessaria per l'esecuzione del docking è ottenuta invece molto più semplicemente: attraverso un comando lanciato da shell si otterrà un file editabile chiamato “`config.txt`” che conterrà tutte le informazioni ed i parametri necessari per creare una griglia adeguata.

Il comando che serve per creare il file di testo è:

- `./pdbbox_vina_15 inut_lig.pdbqt`

Interessante notare che accanto al file di input del ligando intorno al quale deve essere costruita la griglia, viene lanciato il programma eseguibile, fornito insieme a Vina; “`pdbbox_vina_15`”, il quale permette la creazione del file di testo contenete i parametri, come abbiamo visto per AutoDock.

Tale file, che prende automaticamente il nome di “config.txt”, è riportato sotto con dei valori di esempio.

“receptor = viene inserito il file del recettore

ligand = viene inserito il file del ligando su cui effettuare il docking

center_x =
center_y =
center_z =

} coordinate spaziali riguardanti la posizione del ligando cristallografico

size_x =
size_y =
size_z =

} coordinate spaziali riguardanti la grandezza del box formante la griglia

energy_range = indica l’energia massima di differenza che deve esserci tra la migliore binding pose e la peggiore visualizzata dal programma (kcal/mol)”

Una volta editato il file di testo secondo le esigenze lavorative, si può lanciare nella shell di lavoro il comando necessario per l’attuazione del docking molecolare.

- *Vina -config.txt -exhaustiveness 10 -log output.log -out output.pdbqt*

Il primo comando è quello necessario per la specificazione dei parametri e delle coordinate formanti la griglia, mentre il secondo indica il numero di cicli di algoritmo genetico a cui ogni ligando deve essere sottoposto. Esso è importante per determinare l’accuratezza del calcolo. Gli altri comandi invece determinano rispettivamente la creazione di un file di log, utilizzato per controllare la correttezza del processo, e la creazione di un file di output con estensione PDBQT. Al contrario di AutoDock, non si ha bisogno di file di input recanti un’estensione particolare e non si ottengono file di output con un’estensione valida solo per quel determinato programma.

Vina utilizza una particolare scoring function chiamata *united-atom*, che necessita della presenza degli idrogeni polari nelle strutture di input, in modo da poter visualizzare gli atomi pesanti che fungono da donatori di legami ad idrogeno. Tuttavia i gradi di libertà che

muovono gli idrogeni, come i torsionali dei gruppi ossidrilici, sono degenerati. Pertanto, in alcuni file di output, alcuni idrogeni possono essere disposti in modo casuale anche se in accordo con la struttura covalente del ligando.

Nel nostro caso i calcoli sono stati automatizzati utilizzando il seguente script:

```
#!/bin/csh -f

pdbbox_vina_15 ../ADT/lig.pdbqt > log

setenv var `grep "gridcenter" log`

set gct=`echo $var|sed 's/#xyz-coordinates or "auto"/"/|sed 's/REMARK gridcenter //'`

set c1=`echo $gct | awk '{print $1}'`

set c2=`echo $gct | awk '{print $2}'`

set c3=`echo $gct | awk '{print $3}'`

setenv var1 `grep "npts" log`

set gct1=`echo $var1|sed 's/# num.grid points in xyz"/|sed 's/REMARK npts //'`

set d1=`echo $gct1 | awk '{print $1}'`

set d2=`echo $gct1 | awk '{print $2}'`

set d3=`echo $gct1 | awk '{print $3}'`

cat config_or.txt | sed 's/XC/"$c1"/' | sed 's/YC/"$c2"/' | sed 's/ZC/"$c3"/' | sed 's/XS/"$d1"/'
| sed 's/YS/"$d2"/' | sed 's/ZS/"$d3"/' > config_new_or.txt

cat config_new_or.txt | sed 's/,./g' > temp

mv temp config_new_or.txt

echo 'vina --config config_new_or.txt --exhaustiveness 10 --log out.log --out
lig_out.pdbqt'

echo 'pdbqt_to_pdb.py -f lig_out.pdbqt -o lig_out.pdb'

echo 'babel -ipdb lig_out.pdb -h -omol2 ./vina_out.mol2'
```


Affinché questo script vada a buon fine, nella cartella dove si decide di lanciare il docking deve essere presente sia il programma “pdbbox_vina_15” sia il file di parametri chiamato “conFig._or.txt”, modificato secondo le esigenze.

Il template “config._or.txt” indica quali sono i path della proteina e del ligando.

2.1.5 Dock

Il programma DOCK sfrutta una strategia diversa rispetto ai programmi di docking tradizionali, definita *di primo ancoraggio*. Il sito recettoriale viene suddiviso in sfere e il centro di ognuna è considerato come una possibile posizione di un atomo del ligando in esame. Grazie a queste sfere si individua la sottostruttura rigida del composto che viene quindi orientata in diversi modi nel sito recettoriale. Le migliori orientazioni vengono selezionate attraverso una griglia di scoring, che considera il contributo energetico di ognuna di esse. Successivamente il programma costruisce intorno alla sottostruttura le parti flessibili della molecola e, tenendo in considerazione il contesto amminoacidico, può variare leggermente il posizionamento dei ligandi in modo da trovare un minimo energetico. Sfruttando questa strategia DOCK riduce drasticamente il numero di orientazioni possibili in quanto, essendo un programma di docking stocastico, avrebbe un numero enorme di potenziali soluzioni da analizzare.

DOCK fa parte dell’omonima suite di programmi che contiene anche i seguenti software:

- *Sphgen*: identifica il sito attivo e genera le sfere per il docking;
- *Grid*: genera la griglia di scoring

2.3.5.1 Preparazione dei file di input

Per poter effettuare un calcolo con questo programma è necessario attuare alcune procedure di preparazione sia per il ligando da dockare che per la proteina. Per tale scopo è stato utilizzato il software UCSF CHIMERA.

La procedura utilizzata per la preparazione della proteina è la seguente:

- Dalla barra dei menu cliccare *Select/Structure/Ligand* e quindi *Action/Atoms-Bond/Delete* per eliminare l'eventuale ligando cristallizzato;
- Selezionare *Tools/StructureEditing/DockPrep*, si apre una finestra nella quale compare il codice del raggio X e alcune opzioni. Cliccando il pulsante *Ok*, viene visualizzato il riquadro *Add Hydrogens for Dock Prep* dal quale si selezionano i parametri per l'aggiunta degli idrogeni alla proteina;
- Salvare la struttura in formato MOL2;
- Eliminare gli idrogeni dalla proteina e salvare il file in formato PDB assicurandosi di non sovrascriverlo, in quanto è necessario avere a disposizione la struttura con e senza idrogeni.

Il ligando deve essere salvato in formato MOL2. Inoltre deve essere calcolata la carica utilizzando il programma Molcharge. Il lavoro viene iniziato digitando sul terminale la seguente stringa:

- *Molcharge -param molcharge.param*

Dove il file molcharge.param è un file di input così formato:

#Interface settings

#molcharge :

| | |
|------------------------------|------------------------------|
| <i>#-altlocs</i> | <i>false</i> |
| <i>#-aml</i> | <i>false</i> |
| <i>-amlbcc</i> | <i>true</i> |
| <i>#-amberff94</i> | <i>false</i> |
| <i>#-clear</i> | <i>false</i> |
| <i>#-formal</i> | <i>false</i> |
| <i>#-gasteiger</i> | <i>false</i> |
| <i>-in lig.mol2</i> | |
| <i>#-initial</i> | <i>false</i> |
| <i>#-mmff</i> | <i>false</i> |
| <i>#-noh</i> | <i>false</i> |
| <i>#-opls</i> | <i>false</i> |
| <i>-out lig_charged.mol2</i> | |
| <i>-param</i> | <i>molcharge.param</i> |
| <i>#-paramfile</i> | <i>(Not set, no default)</i> |
| <i>#-prefix</i> | <i>molcharge</i> |
| <i>#-singlePoint</i> | <i>false</i> |

Le componenti asterisicate non vengono considerate durante il calcolo, così le uniche variabili rimangono: *-in* = ligando di input, *-amlbcc* = è la funzione utilizzata per assegnare la carica parziale ed aggiungere gli idrogeni, *-out* = nome del file di output, *-param* = nome del file di parametri.

2.3.5.2 Generazione delle sfere e della griglia di scoring

Come precedentemente detto, DOCK necessita di una griglia di scoring e di un file contenente le sfere in cui viene suddiviso il sito recettoriale. È necessario per prima cosa generare la superficie molecolare della proteina, digitando la stringa:

- *Dms proteina_noH.pdb -n -w 1.4 -v -o output.ms*

Dove *dms* è il programma utilizzato, seguito dal file PDB della proteina senza idrogeni, -n = calcola i vettori per i punti della superficie, -w = definisce il raggio della sonda, -v = verboso -o = specifica il nome del file di output.

Si procede editando il file INSPH, contenente i parametri per la generazione delle sfere. È necessario specificare il nome del file di input in formato .ms, appena generato, e quello del file di output (rec.sph).

Il seguente è un esempio del file editato:

```
"rec.ms  
R  
X  
0.0  
4.0  
1.4 rec.sph"
```

Per creare le sfere si procede digitando il comando *sphgen* dal terminale.

Per selezionare solo le sfere che occupano il sito recettoriale è stato lanciato il seguente comando:

- *Sphere_selector rec.sph -lig.mol2 10*

In questo modo vengono selezionate dal file rec.sph solo le sfere che si trovano in un intorno di 10 Å il quale è un valore di default, mentre nel nostro studio sono stati utilizzati valori diversi (vedi paragrafo 3.3) dal ligando cristallografico (lig.mol2). Esse vengono salvate su un file con il nome selected_spheres.sph.

A questo punto si può passare alla generazione del *box*. Procediamo editando il file *box.in*, nel quale va specificato il nome dell'output in formato PDB. Inoltre è necessario modificare

il file in modo che il box venga creato a minimo 5 Å dalla superficie di ciascuna sfera. Il file *box.in* risulta così:

```
"Y
5.0
Selected_spheres.sph
1 box.pdb"
```

Per creare il box attorno alle sfere selezionate si utilizza il comando *showbox box.in*. (vedi paragrafo 3.3). Infine si può procedere con la creazione della griglia sfruttando il programma *Grid*. Esso possiede due differenti funzioni, chiamate *Energy* e *Contact*, che consentono la formazione di due tipi diversi di griglia. In questo caso abbiamo utilizzato la funzione *Energy*, che considera il contributo delle componenti elettrostatiche e di Van der Waals. *Grid* inoltre calcola delle *bumb grid*, cioè delle “griglie di urto” che identificano se un atomo del ligando è sovrapposto al recettore.

Si digita:

- *Gridi -i grid.in -o grid.out*

dove il file *grid.in* contiene i parametri necessari per la generazione della griglia:

| | |
|-------------------------------|-------------|
| <i>compute_grids</i> | <i>yes</i> |
| <i>grid_spacing</i> | <i>0.3</i> |
| <i>output_molecule</i> | <i>no</i> |
| <i>contact_score</i> | <i>no</i> |
| <i>energy_score</i> | <i>yes</i> |
| <i>energy_cutoff_distance</i> | <i>9999</i> |
| <i>atom_model</i> | <i>a</i> |
| <i>attractive_exponent</i> | <i>6</i> |
| <i>repulsive_exponent</i> | <i>12</i> |
| <i>distance_dielectric</i> | <i>yes</i> |

| | |
|----------------------------|--|
| <i>dielectric_factor</i> | <i>4</i> |
| <i>bump_filter</i> | <i>yes</i> |
| <i>bump_overlap</i> | <i>0.75</i> |
| <i>receptor_file</i> | <i>prot_charged.mol2</i> |
| <i>box_file</i> | <i>prot_box.pdb</i> |
| <i>vdw_definition_file</i> | <i>/usr/local/dock6/parameters/vdw_AMBER_parm99.defn</i> |
| <i>score_grid_prefix</i> | <i>grid"</i> |

Vengono generati tre file di output: grid.nrg, grid.bmp e grid.out.

2.3.5.3 Inizio del calcolo di docking

La procedura di docking può essere effettuata considerando il ligando come un'entità rigida o flessibile. Nel corso di questo lavoro di tesi è stata utilizzata la procedura che considera la macromolecola "rigida" ed il ligando "flessibile".

Si può dare inizio al calcolo digitando nella shell il comando:

- *Dock6 -i dock.in -o dock.out*

Dove dock.in è un file di input contenente i seguenti parametri di default (in alcune diciture viene riportato il significato preceduto da asterisco):

| | |
|----------------------------------|--------------------------------|
| <i>"ligand_atom_file</i> | <i>ligando_da_dockare.mol2</i> |
| # ligando su cui fare il docking | |
| <i>limit_max_ligands</i> | <i>no</i> |
| <i>skip_molecule</i> | <i>no</i> |
| <i>read_mol_solvation</i> | <i>no</i> |
| <i>calculate_rmsd</i> | <i>no</i> |
| <i>use_database_filter</i> | <i>no</i> |
| <i>orient_ligand</i> | <i>yes</i> |

| | |
|---------------------------|-------------------------|
| <i>automated_matching</i> | <i>yes</i> |
| <i>receptor_site_file</i> | <i>selected_spheres</i> |

sito del recettore dove prende le sfere

| | |
|-------------------------|-------------|
| <i>max_orientations</i> | <i>1000</i> |
|-------------------------|-------------|

orientazioni massime che può assumere il ligando

| | |
|--------------------------------|------------|
| <i>critical_points</i> | <i>no</i> |
| <i>chemical_matching</i> | <i>no</i> |
| <i>use_ligand_spheres</i> | <i>no</i> |
| <i>use_internal_energy</i> | <i>yes</i> |
| <i>internal_energy_rep_exp</i> | <i>12</i> |
| <i>flexible_ligand</i> | <i>yes</i> |

parametro per tenere il ligando rigido o meno durante il docking

| | |
|------------------------------|-----------|
| <i>user_specified_anchor</i> | <i>no</i> |
| <i>limit_max_anchors</i> | <i>no</i> |
| <i>min_anchor_size</i> | <i>5</i> |

numero minimo di atomi pesanti da considerare per creare un anchor (p.es se il valore è posto 5 un ciclopentano è considerato un anchor mentre un ciclobutano no)

| | |
|-------------------------------|------------|
| <i>pruning_use_clustering</i> | <i>yes</i> |
|-------------------------------|------------|

per clusterizzare in termini energetici le varie pose trovate

| | |
|----------------------------|-------------|
| <i>pruning_max_orients</i> | <i>5000</i> |
|----------------------------|-------------|

massimo numero di orientazioni degli anchor trovati dall'analisi conformazionale

| | |
|----------------------------------|------------|
| <i>pruning_clustering_cutoff</i> | <i>500</i> |
|----------------------------------|------------|

quante molecole considerare dalla ricerca conformazionale in base all'orientazione degli anchor

pruning_conformer_score_cutoff 500

numero massimo di conformeri da prendere dopo minimizzazione

use_clash_overlap *no*

write_growth_tree *no*

bump_filter *yes*

I bumps sono quegli atomi che durante la minimizzazione si trovano sovrapposti agli atomi del recettore e quindi verranno eliminati dal parametro max_bumps

bump_grid_prefix *grid*

max_bumps_anchor 12

numero massimo di atomi che si sovrappongono. Superato questo valore, il filtro bumps viene bypassato

max_bumps_growth 12

score_molecules *yes*

contact_score_primary *no*

contact_score_secondary *no*

grid_score_primary *yes*

grid_score_secondary *no*

grid_score_rep_rad_scale 1

grid_score_vdw_scale 1

grid_score_es_scale 1

grid_score_grid_prefix *grid*

dock3.5_score_secondary *no*

continuous_score_secondary *no*

descriptor_score_secondary *no*

| | |
|--|--|
| <i>gbsa_zou_score_secondary</i> | <i>no</i> |
| <i>gbsa_hawkins_score_secondary</i> | <i>no</i> |
| <i>amber_score_secondary</i> | <i>no</i> |
| <i>minimize_ligand</i> | <i>yes</i> |
| <i>minimize_anchor</i> | <i>yes</i> |
| <i>minimize_flexible_growth</i> | <i>yes</i> |
| <i>use_advanced_simplex_parameters</i> | <i>no</i> |
| <i>simplex_max_cycles</i> | <i>1</i> |
| <i>simplex_score_converge</i> | <i>0.1</i> |
| <i>simplex_cycle_converge</i> | <i>1.0</i> |
| <i>simplex_trans_step</i> | <i>1.0</i> |
| <i>simplex_rot_step</i> | <i>0.1</i> |
| <i>simplex_tors_step</i> | <i>10.0</i> |
| <i>simplex_anchor_max_iterations</i> | <i>500</i> |
| <i>simplex_grow_max_iterations</i> | <i>500</i> |
| <i>simplex_grow_tors_premix_iterations</i> | <i>0</i> |
| <i>simplex_random_seed</i> | <i>0</i> |
| <i>simplex_restraint_min</i> | <i>no</i> |
| <i>atom_model</i> | <i>all</i> |
| <i>vdw_defn_file</i> | <i>/usr/local/dock6/parameters/vdw_AMBER_parm99.defn</i> |
| <i>flex_defn_file</i> | <i>/usr/local/dock6/parameters/flex.defn</i> |
| <i>flex_drive_file</i> | <i>/usr/local/dock6/parameters/flex_drive.tbl</i> |
| <i>ligand_outfile_prefix</i> | <i>output.mol2</i> |

nome del file di output

write_orientations *no*

cluster_conformations *yes*

cluster_rmsd_trheshold *2.0*

#le due funzioni sopra descritte sono state introdotte per effettuare la clusterizzazione delle pose (vedi paragrafo 3.2)

num_scored_conformers *1000*

rank_ligands *no*”

Mentre dock.out è un file di log che contiene varie informazioni riguardo alla procedura di docking appena effettuata.

2.1.6 Fred 3.0

FRED (Fast Rigid Exhaustive Docking) è un programma di docking esaustivo, che quindi deve disporre di una grossa libreria di conformeri derivanti dall’analisi conformazionale dei ligandi. L’analisi conformazionale viene effettuata con il programma **Omega**.

Omega esegue una rapida ricerca di conformeri, ottenendo un rendimento di decine di migliaia di composti al giorno e ciò consente di riprodurre efficacemente le conformazioni bioattive.

Fred, dal canto suo, esamina le possibili pose all’interno del sito attivo della proteina, filtrandole sia per la loro forma complementare, sia per l’analisi delle “features” farmacoforiche. Dopodiché viene scelta una singola pose sulla base del punteggio assegnato dalla scoring function *Chemgauss4*.

Per poter effettuare il docking si deve prima procedere alla formazione di una griglia recettoriale ricorrendo generalmente all’uso dell’interfaccia grafica di Fred, chiamata *Make_receptor*. Va detto che questo è il metodo più intuitivo, ma non è l’unico utilizzabile, poiché è possibile creare un box con pochi semplici comandi lanciati da shell.

Interfaccia grafica: dopo aver caricato sia la proteina sia il ligando originale sulla schermata di Make_receptor, cliccando sulla funzione “Molecular” della sezione “Box” verranno ricercate tutte le possibili tasche recettoriali; quando il processo sarà terminato apparirà la loro superficie colorata in verde. Individuato il sito di interesse, è possibile ottimizzare le forma e le dimensioni del box che lo contiene, che graficamente compare sulla schermata. A questo punto si può procedere con la creazione della griglia recettoriale, in modo da poter analizzare al meglio le interazioni tra molecola e recettore. È possibile utilizzare un filtro di contorno esterno ed un filtro di contorno interno al ligando stesso, che possono rifiutare o meno una pose a seconda della sua complementarità con il sito attivo. Si formeranno due volumi complementari che saranno delimitati da tali contorni. Risulta evidente che la maggior parte degli atomi si adatterà meglio al contorno esterno, anche se per soddisfare tale filtro bisogna che tutti gli atomi pesanti si adattino a questo volume. Per soddisfare il filtro interno basterà invece che solo un atomo pesante si adatti a questo volume. Nel nostro caso è stato attivato il contorno esterno, con un valore di 1000 Å³, mentre è stato disattivato quello interno. Quindi si salva il file contenente i parametri della griglia necessaria per il docking, in formato *oeb*.

Shell: sicuramente un metodo più semplice e veloce, digitando la seguente stringa:

- *receptor_setup -protein prot.mol2 -box lig_or.mol2 -addbox 7.5*

Prima di procedere, come già accennato, si dovrà provvedere all’analisi conformazionale dei ligandi. La stringa utilizzata è la seguente:

- *omega2 -in lig_input.mol2 -out lig_output.mol2 -ewindow 25 -maxconfs 1000 -maxtime 600 -rms 0.2 -strictatomtyping false*

Si noti che accanto ai comandi necessari per creare i file di input e output (“-in” e “-out”) esiste il comando che imposta il massimo valore di energia potenziale (espressa in Kcal/mol) di cui un conformero generato può differire rispetto al minimo calcolato per poter essere salvato (“-ewindow”), il comando che imposta il numero di conformeri che devono essere salvati per ogni singola molecola (“-maxconfs”), quello per impostare il tempo massimo di calcolo tra una molecola e l’altra (espresso in secondi, “-maxtime”) e quello per impostare la distanza cartesiana al di sotto della quale due conformeri verranno scartati in quanto troppo simili (“-rms” = Root Mean Square). Il comando “strictatomtyping” con un settaggio

“true” permette di rifiutare quelle molecole che contengono un atom-type non specificato nel force field MMFF94. Un settaggio “false”, come nel nostro caso, consentirà l’utilizzo dei parametri provenienti da un atom-type simile.

L’analisi conformazionale può essere calcolata anche semplicemente digitando su shell:

- *omega -param omega.parm*

Dove omega.parm è un file di parametri così composto:

#Interface settings

#Execute Options:

#-param (Not set, no default)

#-pvmconf (Not set, no default)

#File Options:

#-commentEnergy false

-in lig.mol2

#-includeInput false

#-log (Not set, no default)

-out lig.oeb.gz

#-param (Not set, no default)

#-pendingFile (Not set, no default)

#-prefix omega2

#-progress none

#-rotorOffsetCompress true

#-sdEnergy false

#-status (Not set, no default)

#-verbose false

#-warts *false*

#3D Construction Parameters:

#-addfraglib *(Not set, no default)*

#-buildff *mmff94s_NoEstat*

#-canonOrder *true*

#-deleteFixHydrogen *true*

#-dielectric *1.000000*

#-exponent *1.000000*

#-fixfile *(Not set, no default)*

#-fixrms *0.150000*

#-fixsmarts *(Not set, no default)*

#-fromCT *true*

#-maxmatch 10

#-setfraglib *(Not set, no default)*

-strictatomtyping *false*

#-strictfrags *false*

#-umatch *true*

#Structure Enumeration:

#-enumNitrogen *true*

#-enumRing *true*

#Torsion Driving Parameters:

#-addtorlib *(Not set, no default)*

#-erange *(Not set, no default)*

| | |
|-----------------------------|------------------------------|
| <i>-ewindow</i> | <i>25.000000</i> |
| <i>#-maxConfRange</i> | <i>(Not set, no default)</i> |
| <i>-maxconfs</i> | <i>400</i> |
| <i>#-maxrot</i> | <i>-1</i> |
| <i>-maxtime</i> | <i>120.000000</i> |
| <i>#-rangeIncrement</i> | <i>5</i> |
| <i>-rms</i> | <i>0.200000</i> |
| <i>#-rmsrange</i> | <i>(Not set, no default)</i> |
| <i>#-searchff</i> | <i>mmff94s_NoEstat</i> |
| <i>#-settorlib</i> | <i>(Not set, no default)</i> |
| <i>#Stereo Parameters:</i> | |
| <i>#-flipper</i> | <i>false</i> |
| <i>#-flipper_maxcenters</i> | <i>12</i> |
| <i>#-strictstereo</i> | <i>true</i> |
| <i>#General:</i> | |
| <i>#-strict</i> | <i>(Not set, no default)</i> |
| <i>#PVM:</i> | |
| <i>#-pvmdebug</i> | <i>false</i> |
| <i>#-pvmlog</i> | <i>(Not set, no default)</i> |
| <i>#-pvmpass</i> | <i>10</i> |

Dato che le frasi precedute dall'asterisco (come già riportato), non vengono contate, i parametri in gioco sono gli stessi (-in, -out, -ewindows, -maxconfs, -maxtime, -rms, strictatomtyping).

Per quanto concerne il docking, conviene prima lanciare un calcolo di prova. Il comando utilizzato è il seguente:

- *fred -rec dock_inp.oeb -dbase lig.mol2 -prefix lig*

dove *-rec* = file di parametri creato con *Make_receptor*, *-dbase* = file contenente la molecola da dockare, e *-prefix* = prefisso dei file di output.

Tra i file di output sarà presente un file contenente tutti i parametri utilizzati per il calcolo, avente l'estensione *.param*:

Parameters settings for FRED 3.0.0

#Interface settings

#Execute Options:

-param fred_settings.param

#-pvmconf (Not set, no default)

#Input:

-receptor receptor.oeb.gz

-dbase lig.oeb.gz

#-confest isomeric

#-molnames (Not set, no default)

#Dock Options:

#-dock_resolution Standard

#Output Files:

-docked_molecule_file docked_fred.sdf

#-undocked_molecule_file undocked.oeb.gz

#-score_file score.txt

#-report_file report.txt

| | |
|--------------------------------|------------------------------|
| <i>#-settings_file</i> | <i>settings.param</i> |
| <i>#-status_file</i> | <i>status.txt</i> |
| <i>#Output Options:</i> | |
| <i>-hitlist_size</i> | <i>1000000</i> |
| <i>-num_poses</i> | <i>1</i> |
| <i>#-annotate_scores</i> | <i>false</i> |
| <i>#-score_tag</i> | <i>(Not set, no default)</i> |
| <i>#-save_component_scores</i> | <i>false</i> |
| <i>#-no_extra_output_files</i> | <i>false</i> |
| <i>#-no_dots</i> | <i>false</i> |
| <i>#-prefix</i> | <i>fred</i> |

Questo potrà essere modificato di volta in volta in base alle esigenze e utilizzato per lanciare il docking con FRED:

- *fred -param docking_fred.param*

2.2 Studi di Dinamica Molecolare

Gli studi di dinamica molecolare sono stati condotti utilizzando il software AMBER14. Questo è in realtà costituito da una suite di programmi che insieme permettono di effettuare simulazioni di dinamica molecolare, specialmente su biomolecole. In realtà nessuno di tali programmi porta questo nome: ad ogni modo, per semplicità, è possibile considerare AMBER come un complesso software che racchiude in sé tutti i vari programmi necessari per la realizzazione di uno studio di dinamica molecolare.

2.2.1 Preparazione Step by Step dei File di Input per la Dinamica Molecolare

Per effettuare uno studio di dinamica molecolare è necessario innanzitutto creare dei file di input, opportunamente elaborati, per tutti gli elementi sottoposti alla simulazione e dei file contenenti i parametri ad essi correlati.

Come da titolo quindi, vediamo step by step le operazioni preliminari da eseguire

Pdb4amber: per prima cosa, omologamente a molti programmi di docking, si deve “pulire” la proteina che sarà oggetto della dinamica, eliminando tutti gli idrogeni e mettendo in ordine tutti i legami di eventuali residui amminoacidici non ben definiti. Questa procedura può essere fatta attraverso interfaccia grafica con il programma *Maestro* o più semplicemente lanciando un'unica riga di codice su shell con il programma *pdb4amber* (programma incluso nella suite di AMBER14):

- `pdb4amber -i prot_input.pdb -o prot_output.pdb --noHyd --dry`

dove `-i` e `-o` indicano rispettivamente i file di input e output, `--noHyd` è il comando per eliminare tutti gli idrogeni e `--dry` è il comando per escludere un'eventuale presenza di acqua. La risoluzione di tutti i residui amminoacidi è invece intrinseca del programma.

Gaussian: Se la simulazione interessa un complesso ligando-recettore, come nel nostro caso, è necessario effettuare il calcolo della carica totale del ligando ma a differenza di quanto descritto sopra e di quanto si faceva solitamente, nel nostro lavoro, questo calcolo non è stato svolto con il classico programma *Antechamber*, ma è stato eseguito con un

programma più preciso e affidabile: *Gaussian*. Partendo dalle leggi fondamentali della meccanica quantistica, Gaussian 09 predice le energie, le strutture molecolari, le frequenze vibrazionali e le proprietà molecolari di ligandi, e reazioni, in un'ampia varietà di ambienti chimici. I modelli di Gaussian 09 possono essere applicati sia a specie stabili che a composti che sono difficili o impossibili da osservare sperimentalmente (per esempio, intermedi di breve durata e strutture di transizione). Ad oggi Gaussian 09 offre le funzionalità di modellazione più avanzate e grazie alla sua ottimizzazione nell'uso della CPU in mono e multiprocessore include tantissimi miglioramenti con cui è possibile modellare sistemi molto complessi anche su hardware modesti. I passaggi da intraprendere possono essere così schematizzati:

- Aprire Gaussian tramite comando su shell `gv`
- Nell'interfaccia grafica, spostare il cursore su "Calculate" e poi su "Gaussian Calculation Setup", su "job type" ed infine selezionare "optimize"
- Cliccare edit e salvare il file con estensione `.com`
- Aprire da editor il file `.com` e sostituire le prime 4 righe con:

```
%nproc=2 #individua il numero di processori
```

```
%mem=1000MB #individua quanta RAM occupare
```

```
%chk=xxx.chk #file di output
```

```
#P opt b3lyp/6-31g(d) geom=connectivity
```

#specifica i dati degli atomi di legame attraverso una sezione di input addizionale e segue la geometria e le modifiche ad esso correlate

- Mandare il calcolo con Gaussian eseguendo su shell il comando `g09 nome_del_file_input_senza_estensione`
- Rinominare il file `.com` come `resp.com` ed anche il file `.chk` come `resp.chk`
- Aprire da editore il file `resp.com` e cancellare tutte le coordinate (dato nel primo calcolo il programma le aveva prese dal file di input, ma adesso dovrà crearle autonomamente secondo le sue previsioni quanto-meccaniche. Oltre a questo sostituire le prime 4 righe del file `resp.com` con:

%nproc=2

%mem=1000MB

%chk=xxx_resp.chk

#P HF/6-31g(d) SCF=Tight #settato di default ed utilizza una combinazione di EDIIS e CDIIS senza damping

geom=check #dal file di input vengono lette solo le cariche

Pop=MK IOp (6/33=2,6/50=1) #produce cariche calcolate secondo il potenziale elettrostatico in punti selezionati secondo lo schema Merz-Singh-Kollman e attraverso il comando IOp si crea il file di output .log utile per gli step successivi

- Mandare il calcolo attraverso la stringa g09 resp e da questo si ottiene il file .log
- Usare il programma antechamber per avere un file finale in .mol2. Quindi su shell mandare la stringa:

antechamber -i resp.log -fi gout -o output.mol2 -fo mol2 -c resp -pf y

- Dato che le coordinate non sono state mantenute, aprire su editore il file .mol2 appena creato e confrontarlo con il ligando di input iniziale (caricato tramite antechamber ma con la funzione bcc), dopo di che sostituire le cariche trovate con il calcolo quanto meccanico con quelle trovate tramite antechamber, negli opportuni atom-type.

Parmchk: a questo punto, con il programma *parmchk* è possibile creare il file contenente i parametri di simulazione relativi al ligando, a partire dall'output di *antechamber*, modificato con le cariche ricavate dal calcolo quanto-meccanico.

Tleap: Una volta creati i parametri relativi al ligando è possibile effettuare l'elaborazione del complesso proteina – ligando attraverso il programma *tleap*. Nelle vecchie release di AMBER veniva utilizzato *xleap*, un programma con la propria interfaccia grafica. Tleap invece viene direttamente eseguito su shell consentendo una creazione di uno script unico, con un'accelerazione notevole in termini di tempo:

tleap -f /usr/local/amber14/dat/leap/cmd/leaprc.ff14SB

source leaprc.gaff

```

fmod = loadamberparams frcmod_lig_charged

fmod2=loadamberparams /usr/local/amber14/dat/leap/parm/frcmod.ionsjc_tip3p

prot = loadpdb ../3OJ8_without_H.pdb

lig = loadmol2 lig_charged.mol2

Set default PBRadii mbondi2

complex = combine {prot lig}

charge complex

addions complex Cl- 0

solvatebox complex TIP3PBOX 20

saveamberparm complex complex_dyn.prmtop complex_dyn.inpcrd

Quit

```

Breve spiegazione dello script: devono essere innanzitutto caricati il force field su cui verranno basati i calcoli di dinamica, il file dei parametri del ligando e quello contenente le coordinate della sola proteina estratta dal complesso. A questo punto deve essere calcolata la carica totale della proteina in modo tale da poterla neutralizzare con l'aggiunta di un numero opportuno di ioni Na⁺ o Cl⁻ a seconda del valore e del segno della carica rilevata. Successivamente è necessario creare un box di acqua che circonda la proteina (in questo caso di raggio 20 Å dalla proteina stessa). Come risultato dell'elaborazione del complesso proteina-ligando vengono infine creati un file contenente le coordinate dell'intero complesso (file.inpcrd) ed uno contenente tutti i relativi parametri (file.prmtop): questi file sono fondamentali per la realizzazione di tutte le simulazioni di dinamica molecolare.

2.2.2 Minimizzazione e Simulazioni di Dinamica

Per effettuare una minimizzazione e/o una simulazione di dinamica molecolare con AMBER14 sono necessari i file .inpcrd e .prmtop di input ed un file contenente i parametri relativi al calcolo.

I parametri di calcolo relativi ad una minimizzazione sono i seguenti:

Imin: permette di scegliere se effettuare la minimizzazione o meno (valore 1 = sì, valore 0 = no). Parametro utilizzato da noi 1

Maxcyc: numero complessivo di cicli di minimizzazione. Parametro utilizzato da noi 5000

Ncyc: numero di cicli di minimizzazione preliminari (effettuati con un force field meno severo). Parametro utilizzato da noi 1000

Ntb: permette di scegliere se effettuare la minimizzazione a volume o a pressione costante (valore 1 = V costante, valore 2 = P costante). Parametro utilizzato da noi 1

Ntr: da attivare se si vogliono impostare dei constraints (valore 1 = sì, valore 0 = no). Parametro utilizzato da noi 1

Cut: permette di impostare la zona entro cui considerare gli atomi non legati. Parametro utilizzato da noi 10

Drms = gradiente di energia in kcal/mol. Se = 1 il suo valore è di -4kcal/mol

Il calcolo della minimizzazione del complesso, viene fatto attraverso la linea di comando ed anche in questo caso può essere creato uno script:

- ```
setenv mp "/usr/local/amber14/bin/mpirun -n (n processori)"

setenv sn "/usr/local/amber14/bin/sander.MPI"

$mp $sn -O -i min1.in -o min1.out -c complesso.inpcrd -p complesso.prmtop -r
complesso.rst -ref complexsso.inpcrd < /dev/null
```

Dove troviamo nelle due prime righe di codice un richiamo al programma *SANDER* per effettuare la minimizzazione, -i e -o come file di input e output, il complesso .inpcrd e prmtop

sono i file creati tramite tleap ed il file .rst è un file di coordinate di output che servirà per la futura dinamica molecolare.

Il comando per mandare la dinamica molecolare è piuttosto semplice e può essere automatizzato sempre tramite script:

- `pmemd.cuda -O -i dyn1.in -o dyn1.out -c complex_min.rst -p complex_dyn.prmtop -r complex_MD1.rst -ref complex_min.rst -x mdcrd_1 < /dev/null'`

Dove `-i` e `-o` sono gli input e gli output dei file di parametri, `-c` e `-ref` fanno riferimento alle coordinate della minimizzazione precedentemente effettuata, `-p` è il complesso su cui fare la dinamica, `-r` è l'output delle coordinate da utilizzare in caso di una dinamica successiva, `-x` sono le coordinate della dinamica.

I parametri relativi alla dinamica molecolare, sono invece più numerosi e complessi rispetto alla minimizzazione e sono raccolti nei file *dyn.in* e possono essere cambiati in base alle proprie esigenze (vedi paragrafo 3.4.1):

**Irest** = permette di riavviare la simulazione da un punto preciso (valore 0 = non si hanno riavvii)

**Ntx** = legge le coordinate, ma non la velocità e le formatta in file ASCII

**Ntpr** = permette di decidere dopo quanti step il programma inizia a scrivere i dati nel file mdinfo

**Ntwr** = permette di decidere dopo quanto il programma fa un salvataggio di un file temporaneo nel caso in cui avvenga un arresto anomalo della dinamica

**Iwrap** = le coordinate della dinamica rimangono simili quelle iniziali

**Ntwx** = ogni “ntwx step” saranno scritte le coordinate nel file della traiettoria (mdcrd)

**Ntr** = permette di fissare gli atomi (specifici) nello spazio

**Nstlim** = numero di MD-step

**Dt** = parametro che imposta ogni quanti femtosecondi viene eseguito uno step

**Gamma\_ln** = frequenza di collisioni delle particelle

**Ntc** = se ha valore 2 la dinamica non considera gli stretching degli H perché sarebbero troppo veloci per essere analizzati

**Ntf** = force evaluation. Valore 1 vengono considerate tutte, valore 2 quelle degli idrogeni vengono scartate, valore 3 nullo

**Ntb** = permette di impostare se eseguire la dinamica a volume costante o pressione costante. Valore uguale a 1 = V costante, valore uguale a 2 = pressione costante

**Cut** = dopo 10 Angstrom non vengono contate le interazioni di non legame

**Pres0** = pressione di riferimento: 1=0.987atm

**Ntt** = algoritmo di elaborazione per l'aumento della temperatura

In base alle condizioni scelte è necessario impostare parametri diversi con informazioni compatibili con la scelta effettuata. Normalmente si mira a realizzare una dinamica molecolare a temperatura e pressione costanti. I primi step di simulazione sono però condotti a basse temperature, in quanto è necessario innanzitutto fornire calore al sistema, e in queste condizioni il calcolo della pressione del sistema è abbastanza inaccurato: la scelta di effettuare questa prima fase di riscaldamento a pressione costante può quindi comportare dei problemi. Per questo motivo la fase iniziale della simulazione, in cui la temperatura del sistema viene fatta passare da 0 K a 300 K, viene condotta a volume costante. Una volta che il sistema si è equilibrato ed ha raggiunto una temperatura stabile intorno a 300 K è possibile passare alle fasi successive di simulazione.

### 2.2.3 Valutazione della Dinamica Molecolare

Effettuata una simulazione di dinamica, risulta ovvio ed utile ai fini della ricerca, valutare i risultati (vedi paragrafo 3.4.1).

I criteri di valutazione sono molteplici: analisi del Root Mean Square Deviation (RMSD), analisi dei valori energetici di Generalized Born Surface Area (GBSA) e Poisson Boltzmann Surface Area (PBSA), analisi di possibili formazioni di nuovi legami a idrogeno.

#### 2.2.3.1 Analisi RMSD

Questa analisi permette di valutare la somiglianza di due pose di una stessa molecola dal punto di vista della disposizione spaziale degli atomi. Ciò può essere fatto grazie al calcolo dell'RMSD (Root Mean Square Deviation), ovvero la distanza quadratica media tra gli atomi delle due pose, misurata in Angstrom (Å). Il valore di RMSD è dato dalla seguente equazione:

$$\text{RMSD} = \sqrt{\frac{1}{N} \sum_{i=1}^N \delta_i^2}$$

N= numero di atomi

δ = distanza tra due atomi corrispondenti

Più il valore di RMSD è piccolo, più due pose sono simili tra di loro. Un valore uguale a 0 equivale infatti a due pose perfettamente identiche. Valori fino a 2 Å indicano comunque una buona somiglianza tra le pose.

Gli studi di RMSD utilizzano un file di parametri rmds\_trajin così composto:

*Trajin* #parametro che importa le traiettorie della dinamica molecolare

*Reference* #coordinate del complesso di riferimento solvatato

*Strip* #comando per rimuovere eventuale acqua o atomi indesiderati (come il sodio o il cloro inclusi nel tleap per azzerare la carica del complesso)



Creato il file `rmsd_trajin`, il calcolo può essere effettuato grazie al programma *cpptraj*, facente parte della suite di AMBER14. Cpptraj è il successore del programma ptraj, utilizzato nelle vecchie release di AMBER. Cpptraj è il programma principale per l'elaborazione dei dati e delle coordinate ricavate dalla dinamica molecolare. È stato sviluppato per essere completamente retro compatibile con gli input presenti nel file di parametri trajin. Cpptraj presenta la stessa velocità di calcolo di ptraj ed in molti casi la sorpassa significativamente, in particolare quando viene svolta l'elaborazione delle traiettorie NetCDF. Inoltre, sono state apportate diverse ottimizzazioni OpenMP, per sfruttare macchine multi-core, sempre in relazione ad una maggiore velocità di calcolo.

Dato che la procedura viene effettuata su shell, si può procedere alla creazione di un semplice script (in alcune diciture viene riportato il significato preceduto da asterisco):

```
cpptraj complex_dyn.prmtop rmsd_trajin
```

```
#analisi dell'RMSD sul complesso .prmtop attraverso i parametri stabiliti nel file
rmsd_trajin
```

```
cat rms_CA.rms | awk '{print $2}' > TMP_prot
```

```
#il file rms_CA.rms riporta i valori di RMSD di ogni singolo step della dinamica dei
carboni in alfa rispetto al complesso di riferimento solvatato e tramite il comando awk
viene racchiusa nel file TMP_prot solo la seconda colonna avente i valori di RMSD (la
prima colonna riporta il numero di step)
```

```
cat rms_LIG.rms | awk '{print $2}' > TMP_lig
```

```
#discorso analogo al file rms_LIG.rms solo che in questo caso i valori riguardano le pose
della dinamica
```

```
grep -v CA_ORIG TMP_prot > TMP
```

```
#la prima riga, CA_ORIG, della colonna presente nel file TMP_lig deve essere eliminata
poiché non è un numero e quindi porterebbe ad errore nel calcolo successivo, ed il file
TMP_prot viene rinominato TMP in modo tale da poter essere letto dall'eseguibile
RMSD_AVER
```

```
./RMSD_AVER
```

```
#è un eseguibile che svolge la media aritmetica del file TMP
```

```
mv AVER AVER_prot
```

#viene rinominato il file di output AVER della media aritmetica, in AVER\_prot

```
rm TMP
```

#rimozione del file TMP

```
grep -v lig_ORIG TMP_lig > TMP
```

#discorso analogo al grep -v CA\_ORIG solo che in questo caso riguarda il ligando

```
./RMSD_AVER
```

```
mv AVER AVER_lig
```

```
rm TMP
```

Il risultato finale riportato nel file AVER\_lig, per quanto riguarda il ligando, e AVER\_prot per quanto riguarda la proteina, sarà quindi una media dell'RMSD di tutti gli step della dinamica (nel nostro caso 200 step) e darà un'approssimazione iniziale su quanto proteina e ligando si siano spostati in termini di coordinate e conformazione rispetto al complesso di riferimento.

Per una valutazione più accurata risulta opportuno importare tutti i dati presenti nei file TMP\_prot e TMP\_lig, in un grafico Excel, ma questo tema verrà affrontato successivamente (vedi paragrafo 3.4.1)

### **2.2.3.2. Analisi MMPBSA**

Questa sezione descrive l'uso dello script python MMPBSA.py per eseguire calcoli dell'Area Superficiale (SA) secondo le equazioni di Poisson Boltzmann (PB) o Generalized Bron (GB). L'analisi dell'MMPBSA è un metodo di post – elaborazione in cui ogni step conformazionale proveniente dalla dinamica viene usato per calcolare le variazioni di energia libera tra due stati (tipicamente uno stato legato ed uno libero). Le differenze di energia libera sono calcolate combinando i cosiddetti apporti energetici della fase gassosa, i quali sono indipendenti dal modello di solvente scelto, con le componenti di energia libera di solvatazione (sia polari che apolari) calcolate da un modello di solvente implicito per ciascuna specie.

Il contributo dell'Entropia, all'energia libera totale, può essere aggiunto per ulteriori affinamenti. I calcoli di entropia possono essere fatti sia in un modello di solvatazione HCT-GB, che in fase gassosa, utilizzando il programma *mmpbsa\_py\_nabnmode* scritto nel linguaggio di programmazione nab.

I contributi dei valori di energia libera della fase gassosa sono calcolate tramite *SANDER*, un programma all'interno della suite AMBER o tramite l'*mmpbsa\_py\_energy*, un programma compreso all'interno del pacchetto “*AmberTools*”, secondo i force field con il quale i vari file sono stati creati. Comunque sia molti tipi comuni di calcoli di MM/PB(GB)SA possono essere eseguiti usando solo il pacchetto AmberTools, mentre alcune delle funzionalità più avanzate richiedono il programma SANDER.

L'energia libera di solvatazione può essere ulteriormente scomposta in un contributo elettrostatico ed in uno idrofobico:

La porzione elettrostatica viene calcolata utilizzando l'equazione di Poisson Boltzmann (PB), il metodo Generalized Born (GB), o il Reference Interaction Model Site (RISM). L'equazione PB è risolta numericamente sia dal programma di PBSA, incluso in AmberTools, che dal programma Adaptive Poisson Boltzmann Solver (APBS) attraverso l'interfaccia iAPBS.

Il contributo idrofobico viene invece approssimativamente calcolato tramite il metodo LCPO implementato in SANDER o tramite il metodo *molsurf* implementato in cpptraj.

L'MM/PB(GB)SA utilizza in genere l'approssimazione che lo spazio configurazionale ha esplorato nei sistemi che sono molto simili tra gli stati legati e non legati, così ogni snapshot per ogni specie viene estratto dallo stesso file di coordinate. Comunque sia lo script MMPBSA.py può accettare file di traiettorie separate per ogni specie. Inoltre, il solvente e gli ioni vengono rimossi dai file delle traiettorie per accelerare la convergenza impedendo interazioni solvente – solvente.

Vediamo adesso nel dettaglio come eseguire i calcoli di GBSA e PBSA utilizzando il programma MMPBSA.py incluso nel pacchetto AmberTools:

**Tleap** = come descritto per la dinamica molecolare, anche in questo caso dovremmo preparare dei file di input. Anche se il ligando ed il recettore sono gli stessi, non possono essere usati gli stessi file di input della dinamica molecolare, poiché questi sono

complessati con l'acqua ed in un calcolo MMPBSA l'acqua non deve essere presente. Di seguito viene riportato lo script:

```
tleap -f /usr/local/amber14/dat/leap/cmd/leaprc.ff14SB

source leaprc.gaff

source leaprc.ff14SB

setenv AMBERHOME /usr/local/amber14/

prot = loadpdb 3OJ8_without_H.pdb

fmod = loadamberparams frcmod_lig_charged

lig = loadmol2 lig_charged.mol2

set default PBRadii mbondi2

saveamberparm prot prot.prmtop prot.inpcrd

saveamberparm lig lig.prmtop lig.inpcrd

complex = combine {prot lig}

saveamberparm complex complex_noWater.prmtop complex_noWater.inpcrd
```

Come si può vedere lo script tleap risulta identico a quello utilizzato per creare i file di input della dinamica molecolare, eccetto il fatto che in questo caso al complesso non viene aggiunta acqua ed al fatto che viene aggiunto un parametro in più “Set default PBRadii mbondi2” il quale viene utilizzato per scegliere i set di raggi atomici da utilizzare per i calcoli di GBSA e PBSA (di seguito verrà descritta l'importanza di questo valore).

**MMPBSA.py:** creati i file di input, allo stesso modo della dinamica, il calcolo viene fatto attraverso una riga di comando:

- *\$AMBERHOME/bin/mpirun -np 4 MMPBSA.py.MPI -O -i mmpbsa.in -o FINAL\_RES.dat -sp complex\_dyn.prmtop -cp complex\_noWater.prmtop -rp prot.prmtop -lp lig.prmtop -y mdcrd\_?*

Dove `-np` = numero di processori utilizzati, `-i` e `-o` sono i file di input e output, `-sp` = complesso proteina–ligando solvatato, `-cp` = complesso non solvatato, `-rp` = recettore, `-lp` = ligando e `-y` = input delle traiettorie da analizzare.

Il calcolo si appoggia ad un file di parametri nominato *mmpbsa.in* dove troviamo le seguenti diciture:

**Keep file** = parametro che permette di decidere se salvare o meno i file temporanei. Utile nel caso di un arresto anomalo del calcolo (valore = 2 salva tutto, valore = 1 non salva niente)

**Igb** = parametro che permette di decidere quale Generalized Bond utilizzare. Questo parametro è strettamente correlato al parametro `PBRadii mbondi` prima descritto. Infatti se l'igb ha un valore = 2 o 5 il comando `PBRadii mbondi` deve essere 2, se invece l'igb ha un valore = 8 deve essere utilizzato il comando `PBRadii mbondi3`. Nel nostro caso abbiamo utilizzato un valore 2 ed un `mbondi2`.

**Saltcon** = concentrazione del sale in molarità (default = 0.000. Nostro calcolo = 0.150)

**Istrng** = forza ionica espressa in Molarità

**Startframe** = parametro che permette di decidere da quale step della dinamica far iniziare il calcolo (default = 51)

**Radiopt** = questo è un valore correlato al `PBRadii` e di conseguenza all'igb. Infatti questo parametro permette di settare il `PBRadii` da utilizzare e se viene impostato = 0 utilizza il `PBRadii` impostato nel file `complex_dyn.prmtop`

**Inp** = nuovo parametro della suite AMBER14 che permette di non calcolare l'energia di dispersione della solvatazione ma solo quella all'interno della cavità. Se il suo valore è impostato a 2 viene eseguito un calcolo automatico di entrambe le energie (default = 2).

### 2.2.3.3 Analisi di Nuovi Legami a Idrogeno

Nel nostro caso, il ligando vedeva solo interazioni idrofobiche (vedi paragrafo 3.4.1). Qui si riporta solo il comando ed il file di parametri per eseguire il calcolo attraverso il già citato programma cpptraj (vedi paragrafo 2.4.3):

- *cpptraj complex.prmtop Hbond\_ale*

Come nel caso del calcolo dell'RMSD il programma cpptraj prevede una sola riga di comando dove si trova il complesso da analizzare ed il file di parametri Hbond\_ale, contenente gli input necessari per avviare la sequenza:

*Trajin #coordinate del complesso da analizzare*

*Hbond #in questo parametro si può decidere quali amminoacidi definire come donatori di legame a idrogeno e quali atomi del ligando definire come accettori di legame a idrogeno*

*.dat #file di output*

## 2.3 Creazione di un File Average per Ulteriori Studi Computazionali

Come facilmente intuibile, il risultato finale di una dinamica molecolare è composto da una serie di frame racchiusi in un unico file di output. Risulta evidente dunque la necessità di convergere tutti questi frame attraverso una media aritmetica, delle coordinate spaziali e conformazionali, in modo da aver un risultato finale leggibile e riproducibile per ulteriori test di docking e di analisi farmacoforiche.

A questo scopo viene nuovamente utilizzato il programma cpptraj con le opportune modifiche del file di parametri AVER:

*trajin #coordinate del complesso di cui fare la media con gli step di mio interesse*

*average output.rst #viene inserito il nome dell'output in formato rst perché dopo verrà fatta una minimizzazione (vedi step successivo)*

:1-543 #vengono inseriti gli atomi di interesse per il calcolo della media (nel nostro caso tutti, proteina più ligando (senza acqua)

Lo step successivo sarà una minimizzazione secondo le procedure descritte nel paragrafo 2.4.2, dato che il risultato finale dell'average è descritto come un calcolo puramente matematico, poco compatibile con le future simulazioni biologiche.

Di seguito viene riportato uno script creato dal sottoscritto, per automatizzare i calcoli di average, minimizzazione e calcolo RMSD tra il composto finale e gli step della dinamica (in alcune diciture viene riportato il significato preceduto da asterisco):

```
#!/bin/csh -f

echo 'setenv dir `pwd`' >> ./script_average

foreach file (/home/alex/nfs/MD/lig_charged/*.mol2)

set code=`echo $file|sed 's/.mol2//'|sed 's:/home/alex/nfs/MD/lig_charged/::'`

echo 'cd $dir/$code'/AVER_f*' >> ./script_average

echo 'cpptraj ../$code'_dyn.prmtop aver' >> ./script_average

#calcolo dell'average

echo 'setenv mp "/usr/local/amber14/bin/mpirun -n 6"' >> ./script_average

echo 'setenv sn "/usr/local/amber14/bin/sander.MPI"' >> ./script_average

echo '$mp $sn -O -i ./min_av.in -o min_av.out -c complex_average_cpptraj.rst -p
../$code'_noWater.prmtop -r complex_average_min.rst -ref
complex_average_cpptraj.rst < /dev/null' >> ./script_average

#calcolo della minimizzazione

echo 'ambpdb -p ../$code'_noWater.prmtop <complex_average_min.rst>
complex_average_min.pdb' >> ./script_average

#conversione del file della minimizzazione in formato .pdb

echo 'cpptraj ../$code'_dyn.prmtop rmsd_trajin_aver_noW' >> ./script_average
```

#formazione delle coordinate di output senza acqua e Cloro per il seguente calcolo di RMSD:

```
echo 'cpptraj ../$code'_noWater.prmtop rmsd_trajin_aver' >> ./script_average
```

#calcolo dell'RMSD tra il composto della dinamica non solvatato e l'average

```
echo 'cat rms_CA.rms | awk "{print $2}" > TMP_prot' >> ./script_average
```

```
echo 'cat rms_LIG.rms | awk "{print $2}" > TMP_lig' >> ./script_average
```

```
echo 'grep -v CA_ORIG TMP_prot > TMP' >> ./script_average
```

```
echo './RMSD_AVER' >> ./script_average
```

```
echo 'mv AVER AVER_prot' >> ./script_average
```

```
echo 'rm TMP TMP_prot' >> ./script_average
```

```
echo 'grep -v lig_ORIG TMP_lig > TMP' >> ./script_average
```

```
echo './RMSD_AVER' >> ./script_average
```

```
echo 'mv AVER AVER_lig' >> ./script_average
```

```
echo 'rm TMP TMP_lig' >> ./script_average
```

```
echo cd ../
```

```
end
```



## CAPITOLO 3 PARTE SPERIMENTALE

Lo scopo di questa tesi è stato quello di identificare il corretto binding mode del composto VS8 (individuata tramite studi precedenti svolti nel nostro laboratorio [147]), nel sito catalitico della FAAH, requisito fondamentale per lo sviluppo e l'ottimizzazione di questa molecola. La prima parte del progetto si è incentrata sullo sviluppo di un robusto protocollo di docking che consentisse di campionare tutte le possibili disposizioni spaziali del composto VS8, all'interno del sito catalitico della FAAH, definendo la possibile conformazione attiva. Infine, la stabilità delle interazioni è stata verificata attraverso studi di dinamica molecolare.

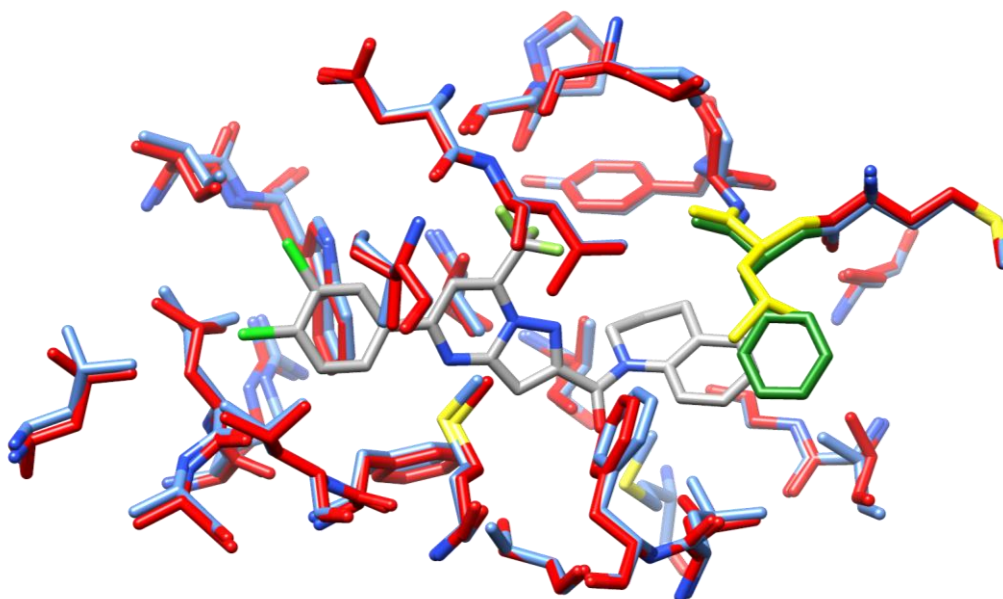
La conformazione bioattiva individuata al termine dell'intero protocollo sarà quindi presa come riferimento per studi di lead optimization mirati allo sviluppo di analoghi con migliore attività inibitoria.

### *3.1 Analisi delle Strutture Cristallografiche*

Come già riportato al paragrafo 1.4, in letteratura esistono soltanto due cristalli della FAAH complessata con un inibitore non covalente. Tali cristalli, appartenenti alla proteina di ratto, sono stati scaricati dal Protein Data Bank (codici PDB: 3QK5 e 3QJ9) ed utilizzati per l'analisi del sito attivo della FAAH. Tuttavia le differenze tra la r/FAAH e la h/FAAH (vedi paragrafo 1.2.1) non dovrebbero essere trascurate se l'obiettivo primario è quello di sviluppare farmaci per uso umano, infatti anche se le due proteine sono caratterizzate dalla stessa triade catalitica Ser241.Ser217-Lys142, condividono omologie amminoacidiche per circa l'82% presentando nel sito attivo sei mutazioni, che passando dal ratto all'umano sono: Leu192Phe, Phe194Tyr, Ala377Thr, Ser435Asn, Ile491Val, Val495Met. È stato quindi deciso di prendere come riferimento per le procedure di questo lavoro di tesi la struttura cristallografica della h/r FAAH (codice PDB: 3OJ8).

In letteratura sono presenti attualmente ben 13 cristalli della h/r FAAH, molti dei quali rappresentano complessi con inibitori  $\alpha$ -chetoeterociclici covalenti. Analizzando attentamente le interazioni ligando-r/FAAH nel complesso 3QJ9, *Min e colleghi* [24] hanno notato come la mutazione Leu192Phe sia l'unica ad avere un impatto potenzialmente

rilevante sull'affinità di tale composto passando dalla proteina di ratto a quella umana. La Phe192 apparirebbe in grado di stabilire un'importante interazione  $\pi$ - $\pi$  stacking con la porzione fenilica dell'anello benzoimidazolico, giustificando la maggiore attività del composto sulla h/FAAH rispetto all'enzima di ratto ( $IC_{50} = 2 \text{ nM}$  vs  $18 \text{ nM}$ ). Le stesse considerazioni possono essere fatte per il complesso 3QK5. È stato quindi deciso di scegliere la struttura della h/r FAAH che presentasse il maggior grado di similarità tra la conformazione della Phe192 e la Leu192 della r/FAAH (Fig.19).



*Fig.19 Sito catalitico della h/FAAH (in celeste), sovrapposta alla r/FAAH di ratto (in rosso). È evidenziata la mutazione Phe192 (in verde) Leu (in giallo) tra l'enzima umano e quello di ratto. In grigio la molecola VS8.*

### ***3.2 Studio Preliminare di Consensus Docking sulla Pose VS8***

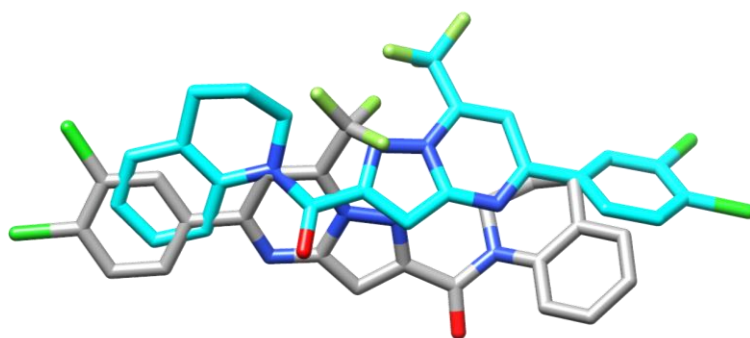
Come strumento per verificare la disposizione spaziale del composto VS8 all'interno del sito catalitico dell'enzima FAAH individuata dallo studio sopra citato, è stato eseguito uno protocollo di consensus docking. Con questa procedura ogni molecola viene processata da più programmi di docking, fornendo per ognuna di esse varie pose all'interno del sito di legame. Queste vengono confrontate tra di loro e solo le pose che vengono predette da un

numero elevato di metodi sono considerate affidabili. Studi effettuati nel nostro laboratorio hanno infatti dimostrato come, più alto è il numero di procedure di docking che predicono una determinata pose, più tale pose si avvicina a quella reale. Per questo motivo un approccio di tipo consensus può essere utilizzato come indice di affidabilità di una pose derivante da studi di docking. In questo modo vengono scartati numerosi falsi positivi, poiché anche se un programma ci fornisse un risultato erraneo è molto probabile che gli altri lo scartino, e quindi che complessivamente l'errore venga corretto. Per questo studio è stato deciso di ricorrere a 6 programmi di docking, per un totale di 10 metodi: GOLD, con le sue quattro funzioni di scoring (asp, plp, chemscore, goldscore), Glide, nelle modalità SP e XP, AutoDock, Vina, DOCK e FRED.

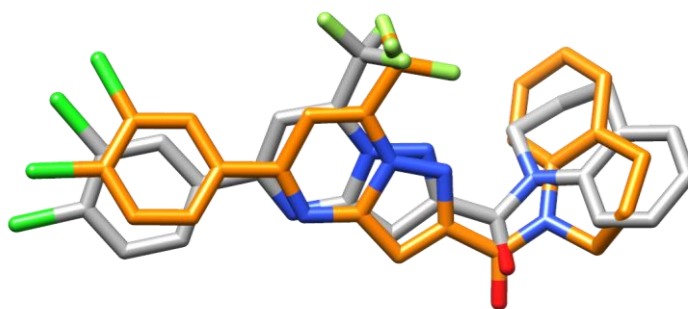
I calcoli sono stati effettuati secondo le modalità descritte al paragrafo 2.1, apportando alcune modifiche agli input di default, per standardizzare, e quindi migliorare, le procedure. Come descritto al paragrafo 2.2.3.1, per valutare quali sono le pose predette da un numero elevato di metodi è necessario ordinarle sulla base della reciproca distanza tra l'una e l'altra. Questa operazione prende il nome di *clusterizzazione* e consente di raggruppare le pose sulla base di un valore soglia di RMSD scelto opportunamente. Ad esempio, se per un certo valore soglia di RMSD le dieci pose sono raggruppate in uno o più cluster (gruppi), all'interno di ciascun cluster ci saranno tutte le pose che tra loro hanno un RMSD sempre inferiore al valore soglia stabilito. Nel nostro caso è stato scelto un valore soglia di 2 Å, che garantisce l'elevata somiglianza delle pose all'interno di ogni singolo cluster.

Una prima analisi svolta sul composto VS8 è stata eseguita senza apportare modifiche ai valori standard dei vari programmi di docking, con un obiettivo di indagine preliminare.

È stata quindi effettuata un'analisi conformazionale della pose e sono state messe in atto le 10 procedure di docking. Come prevedibile il risultato del consensus è stato molto alto, 8/10, dove solo il programma FRED e GOLD, con la funzione GoldScore, hanno riportato una disposizione della conformazione VS8 diversa rispetto a quella di partenza (Fig. 20) (Fig. 21).



*Fig. 20 Sovrapposizione della pose VS8 (grigio) alla pose individuata tramite il programma GOLD con la funzione GoldScore*



*Fig. 21 Sovrapposizione della pose VS8 (grigio), alla pose individuata tramite il programma FRED*

### ***3.3 Analisi di Docking Approfondite***

Per aprire la strada verso l'individuazione del giusto binding mode del composto VS8 all'interno della FAAH, sono state effettuate procedure di docking robusto, destinate ad ampliare il ventaglio di possibili conformazioni che la molecola potrebbe assumere nel sito catalitico.

A questo scopo l'analisi di docking ha previsto come obiettivo la creazione di diversi cluster all'interno di ogni programma, formati attraverso il ranking di RMSD, binding mode ed energia di legame proteina–ligando, e solo i 5 più popolati sono stati tenuti in considerazione e sono stati confrontati. Questa procedura però, non è stata possibile attuarla su tutti i metodi di docking, ma solo su quelli che permettevano una modifica dei parametri di input, tra cui: AutoDock, Dock e le 4 funzioni di GOLD.

Per standardizzare ulteriormente i dati di input, riuscendo così ad ottenere risultati più affidabili e conformi, attraverso l'interfaccia grafica di AutoDock è stato costruito un box del sito attivo, le cui misure e coordinate tridimensionali contenute nel file prot.gpf sono state usate anche nel programma Dock.

*Dimensioni:*

$x = 67$

$y = 74$

$z = 55$

*Coordinate*

$x = -15.832$

$y = -0.5$

$z = -5.263$

Quindi i calcoli di docking sono stati effettuati con GOLD secondo le procedure viste al paragrafo 2.1.1, con AutoDock secondo le procedure previste al paragrafo 2.1.3, mentre per Dock è stata necessaria una modifica della procedura descritta al paragrafo 2.1.5, a causa dell'uniformità del box rispetto ad AutoDock:

**Generazione delle sfere:** nel nostro caso il raggio atomico delle sfere generate dal comando sphere\_selector è stato di 7 Å rispetto ai 10 Å di default. Inoltre siccome le sfere trovate dal programma di docking risultavano mal distribuite all'interno della

cavità di legame, è stata necessaria una modifica di queste attraverso l'interfaccia grafica CHIMERA, dove seguendo la sequenza actions – label – altLoc – sphgenNi è stato possibile vedere il numero relativo di ogni sfera, in modo tale da procedere ad una loro eliminazione, con editore, dal file *selected\_sphere*.

**Box manuale:** la procedura che prevede la creazione di un box secondo i propri standard si attua tramite il comando domanda/risposta showbox:

*Showbox*

*Automatically construct box to enclose spheres? N* #con la funzione N (No), si indica che la creazione del box avverrà manualmente

*Box center user-defined or sphere cluster center of mass? U* #con la funzione U (user-defined) si indica che il box non sarà costruito intorno al file *selected\_sphere.sph*, ma sarà costruito attraverso le coordinate e dimensioni da noi impostate

*Enter box center coordinates -15.832 -1.5 -5.263* #in questo input verranno inserite le coordinate del box modificato tramite AutoDock (x y z)

*Enter box dimension 35 50 35* #attraverso questo input verrà impostata la dimensione del box, ma in questo caso i dati copiati dal file *prot.gpf* di AutoDock, dovranno essere moltiplicati per 0.375 in modo da ottenere misure in Amstrong

*Output filename*

### **3.3.1 Analisi dei risultati di docking**

Effettuati gli studi di docking è stata necessaria una loro attenta valutazione, per scremare le soluzioni trovate. Infatti, attraverso i parametri da noi stabiliti i risultati di docking per ogni programma sono stati:

**GOLD** = tutte e 4 le funzioni utilizzate hanno individuato 1000 pose racchiuse in 5 cluster (vedi file *.conf* paragrafo 2.1.1);

**Dock** = il programma ha individuato 1702 pose racchiuse in 238 cluster;

**AutoDock** = il programma ha riportato 200 pose racchiuse in 8 cluster.

Come suggerito nel paragrafo 3.3, per poter affrontare gli step successivi, il numero di pose da analizzare non poteva essere troppo elevato, quindi attraverso l'interfaccia grafica di Maestro sono stati confrontati e selezionati solo i rappresentanti dei migliori 5 cluster di ogni programma (30 pose) sulla base di:

- Energia di legame;
- Popolazione del Cluster ;
- Somiglianze strutturali: nel caso in cui 2 pose avessero avuto ottima energia di legame e ottima popolazione del cluster, ma presentavano troppe somiglianze strutturali, ne veniva scelta solo una.

Come riportato in tabella 1, solo 15 tra le pose analizzate sono state prese in considerazione e quindi ulteriormente studiate mediante analisi di dinamica molecolare.

| List Cluster | Docking       | N° cluster | Population | Energy (Kcal/mol) |
|--------------|---------------|------------|------------|-------------------|
| 1            | DOCK_10       | 1/238      | 19/1702    | -38.50            |
| 2            | GOLD_ASP_1    | 1/5        | 991/1000   | <i>nd</i>         |
| 3            | GOLD_GSCORE_1 | 1/5        | 161/1000   | <i>nd</i>         |
| 4            | DOCK_4        | 32/238     | 29/1702    | -16.45            |
| 5            | ADT_3         | 3/8        | 22/200     | -10.4             |
| 6            | DOCK_7        | 5/238      | 26/1702    | -32.48            |
| 7            | DOCK_3        | 4/238      | 29/1702    | -35.67            |
| 8            | DOCK_2        | 2/238      | 32/1702    | -37.73            |
| 9            | DOCK_5        | 56/238     | 28/1702    | -3.31             |
| 10           | DOCK_8        | 8/238      | 20/1702    | -29.65            |
| 11           | GOLD_ASP_3    | 3/5        | 3/1000     | <i>nd</i>         |
| 12           | GOLD_GSCORE_2 | 2/5        | 454/1000   | <i>nd</i>         |
| 13           | GOLD_PLP_1    | 1/5        | 104/1000   | <i>nd</i>         |
| 14           | GOLD_GSCORE_3 | 3/5        | 183/1000   | <i>nd</i>         |
| 15           | DOCK_11       | 52/238     | 17/1702    | -5.35             |

*Tabella 1. Pose individuate e portate allo step successivo*

### ***3.4 Studi di Dinamica Molecolare***

Nell'ultima fase di ricerca è stato realizzato uno studio di dinamica molecolare, grazie al software AMBER14 (Assisted Model Building with Energy Refinement), descritto al paragrafo 2.2. Attraverso la dinamica molecolare è possibile verificare la stabilità di un sistema biologico fornendogli energia. Nel nostro caso, uno studio di questo tipo è stato condotto per verificare la stabilità delle interazioni dei complessi FAAH-ligando realizzati attraverso i calcoli di docking. Al contrario del docking, in cui si analizzano modelli recettoriali statici, nella dinamica anche la proteina acquisisce una certa capacità di movimento e si ha quindi una riproduzione più fedele delle condizioni biologiche. Se le interazioni stabilite non sono sufficientemente forti, il complesso potrà andare incontro a dissociazione.

#### ***3.4.1 Simulazioni di dinamica molecolare e analisi dei risultati***

Sono stati sottoposti a dinamica molecolare tutti e 15 i complessi che hanno superato il filtro di clusterizzazione. La creazione dei parametri e dei file di input necessari ad AMBER per effettuare i calcoli è stata realizzata attraverso una procedura automatizzata. È stato deciso di eseguire per ogni complesso lo stesso protocollo di 20ns di simulazione (vedi paragrafo 2.2.2), suddivisi rispettivamente in:

**0.5 ns** in cui il sistema porta la temperatura da 0 K a 300 K a volume costante, tenendo tutti i carboni in alfa bloccati;

**3 ns** necessari per la stabilizzazione del sistema, effettuati a pressione costante, sempre con i carboni in alfa bloccati;

**16.5 ns** di dinamica libera, a pressione costante, in cui i carboni in alfa non sono bloccati.

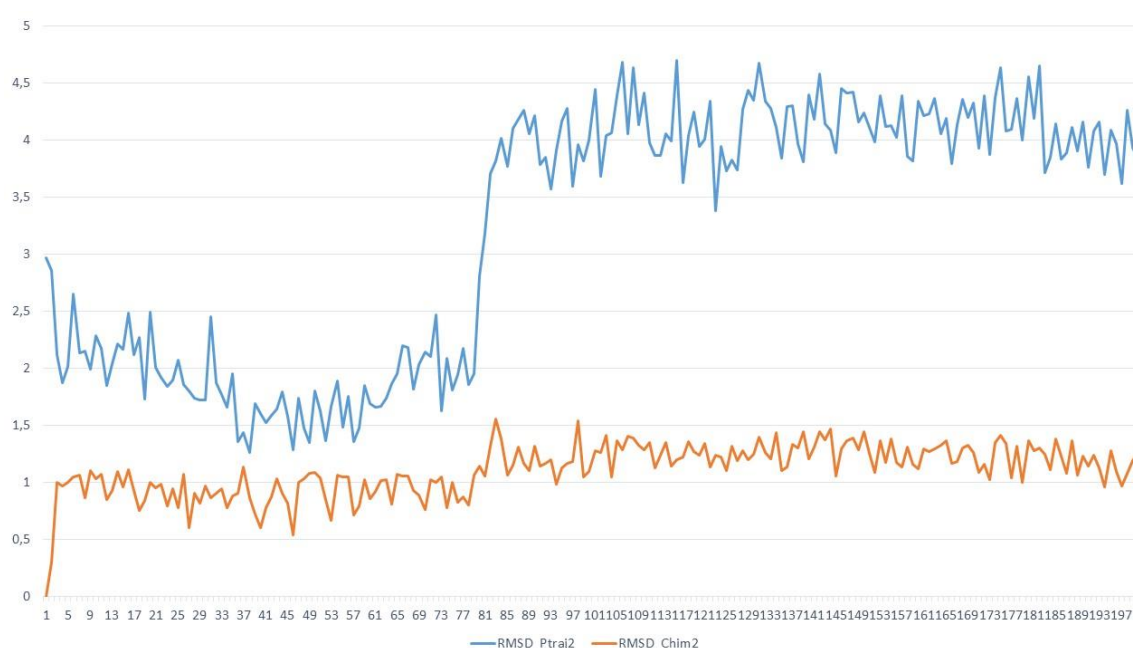
I risultati sono stati analizzati sulla base delle metodiche riportate al paragrafo 2.2.3.1, 2.2.3.2 e 2.2.3.3, e dato che lo scopo oggetto di questo studio è stato quello di trovare il più probabile binding mode del composto VS8, è risultato necessario e conveniente riportare su tabella tutti i valori trovati per ogni singola dinamica molecolare (vedi Tab. 2):



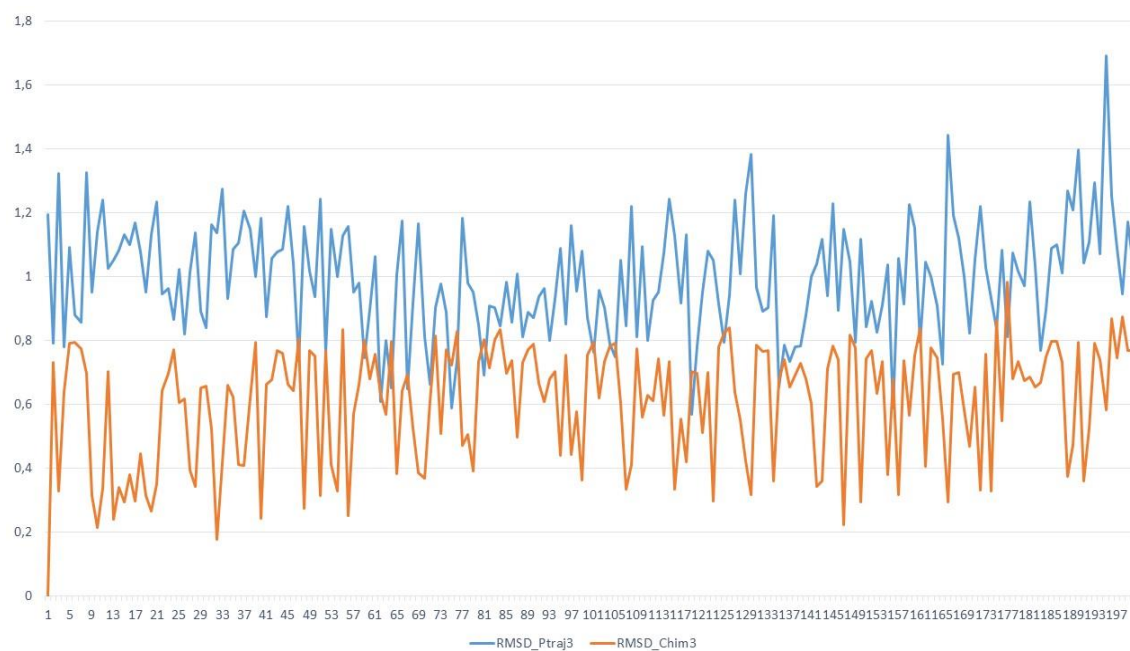
| Pose | MMPBS<br>A<br>(Kcal/mol) | RM<br>SD_<br>Prot<br>(Å) | RM<br>SD_<br>Lig<br>(Å) | H-<br>Bond<br>% | Docking program | N°<br>cluster | Populati<br>on |
|------|--------------------------|--------------------------|-------------------------|-----------------|-----------------|---------------|----------------|
| 1    | ΔGB -38<br>ΔPB -32       | 0.84                     | 3.96                    | 1%              | DOCK_10         | 1/238         | 19/1702        |
| 2    | ΔGB -49<br>ΔPB -38       | 0.84                     | 3.22                    | 0.5%            | GOLD_ASP_1      | 1/5           | 991/1000       |
| 3    | ΔGB -52<br>ΔPB -30       | 0.83                     | 0.99                    | 25%             | GOLD_GSCORE_1   | 1/5           | 161/1000       |
| 4    | ΔGB -48<br>ΔPB -33       | 0.89                     | 1.32                    | 6%              | DOCK_4          | 32/238        | 29/1702        |
| 5    | ΔGB -41<br>ΔPB -33       | 0.80                     | 2.54                    | 56%             | ADT_3           | 3/8           | 22/200         |
| 6    | ΔGB -42<br>ΔPB -36       | 0.98                     | 2.90                    | 8%              | DOCK_7          | 5/238         | 26/1702        |
| 7    | ΔGB -41<br>ΔPB -35       | 0.85                     | 2.86                    |                 | DOCK_3          | 4/238         | 29/1702        |
| 8    | ΔGB -35<br>ΔPB -32       | 0.91                     | 2.62                    | 0%              | DOCK_2          | 2/238         | 32/1702        |
| 9    | ΔGB -50<br>ΔPB -38       | 0.88                     | 1.49                    | 10%             | DOCK_5          | 56/238        | 28/1702        |
| 10   | ΔGB -40<br>ΔPB -36       | 0.80                     | 2.79                    | 10%             | DOCK_8          | 8/238         | 20/1702        |
| 11   | ΔGB -49<br>ΔPB -37       | 0.89                     | 1.38                    | 0.6%            | GOLD_ASP_3      | 3/5           | 3/1000         |
| 12   | ΔGB -45<br>ΔPB -34       | 0.84                     | 2.55                    | 14%             | GOLD_GSCORE_2   | 2/5           | 454/1000       |
| 13   | ΔGB -48<br>ΔPB -36       | 0.97                     | 2.38                    | 8%              | GOLD_PLP_1      | 1/5           | 104/1000       |
| 14   | ΔGB -52<br>ΔPB -36       | 0.79                     | 3.16                    | 0%              | GOLD_GSCORE_3   | 3/5           | 183/1000       |
| 15   | ΔGB -47<br>ΔPB -32       | 0.85                     | 1.53                    | 0.5%            | DOCK_11         | 52/238        | 17/1702        |

Tab. 2 Tabella riassuntiva dei risultati ottenuti dall'analisi della dinamica molecolare

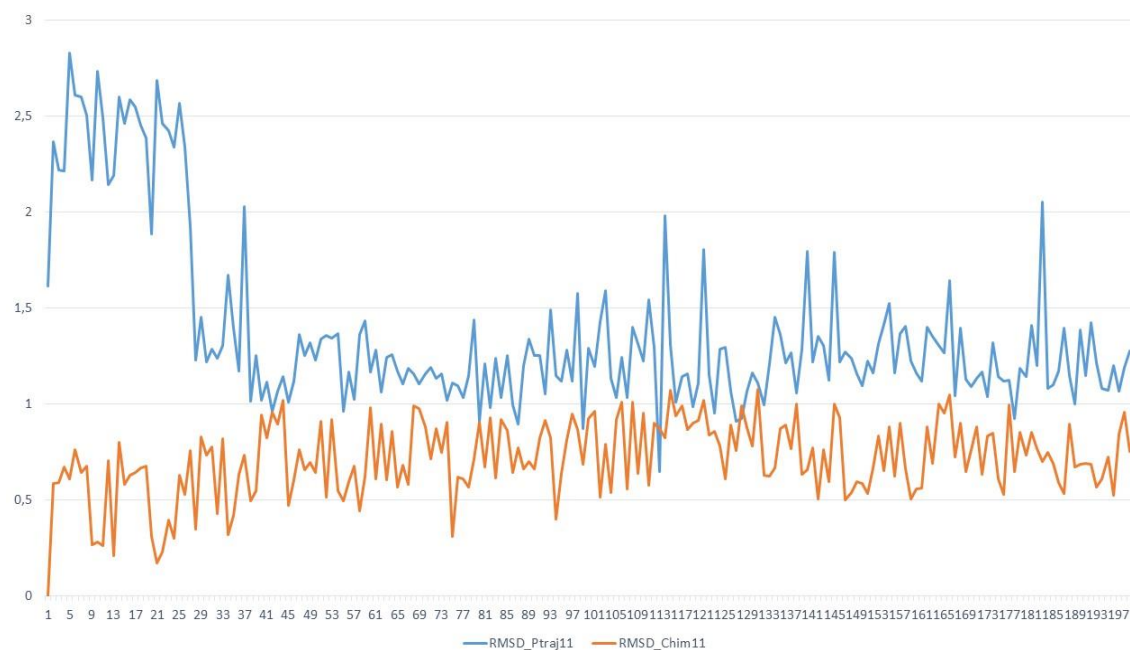
I valori di RMSD riportati in tabella 2, sono valori di RMSD medio, valori che possono permettere una prima analisi sulla bontà della dinamica, ma che non possono portare a considerazioni finali. A questo scopo è risultato utile portare su grafico i dati RMSD di ogni frame di ogni dinamica, per confrontare la reale stabilità della pose e dei carboni in alfa della proteina durante le alterazioni energetiche. Di seguito sono riportati i grafici delle dinamiche che hanno avuto maggior successo in relazione a tutti i dati finora forniti:



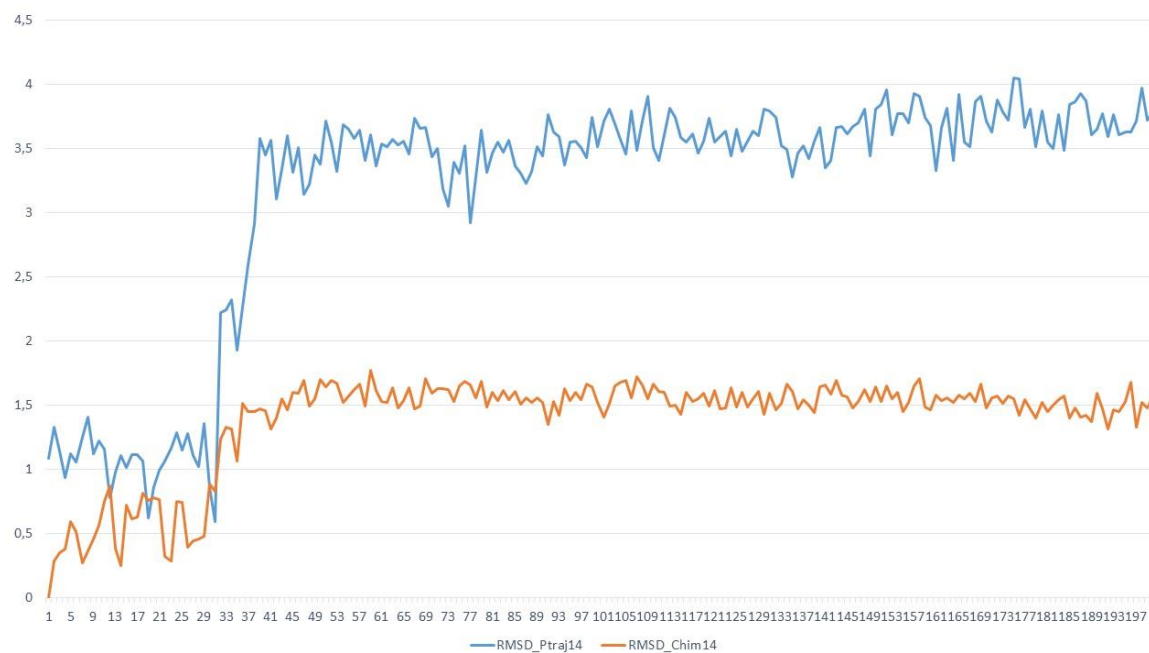
*Fig. 22 Pose 2: andamento dell'RMSD della pose (blu), andamento dell'RMSD dei carboni in alfa della proteina (arancio)*



*Fig. 23 Pose 3: andamento dell'RMSD del pose (blu), andamento dell'RMSD dei carboni in alfa della proteina (arancio)*



*Fig. 24 Pose 11: andamento dell'RMSD del pose (blu), andamento dell'RMSD dei carboni in alfa della proteina (arancio)*

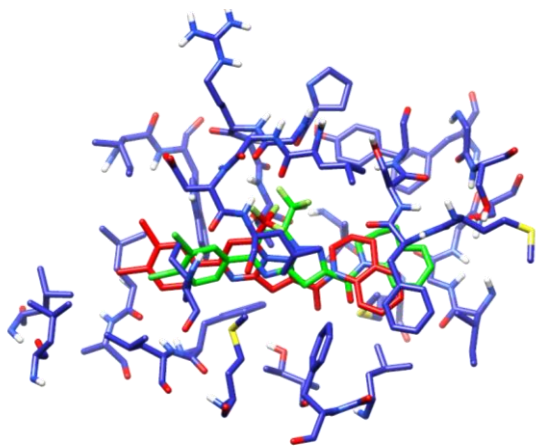


*Fig. 25 Pose 14: andamento dell'RMSD della pose (blu), andamento dell'RMSD dei carboni in alfa della proteina (arancio)*

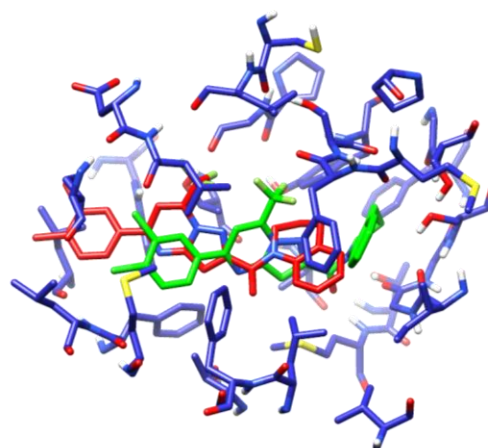
Come facilmente dimostrabile dall'analisi dei grafici appena riportati il valore dell'RMSD dei carboni in alfa della proteina risulta costante in tutte e 4 le dinamiche.

Analizzando le varie pose invece:

- Fig. 22 : l'andamento del grafico può essere suddiviso in 2 porzioni ben distinte. La spiegazione di questo fenomeno consiste nel fatto che la pose assume 2 conformazioni diverse, la cui seconda (Fig. 27), nonostante trovi un RMSD maggiore rispetto alla prima (Fig 26), risulta più stabile e favorita con un RMSD medio di 0.96;

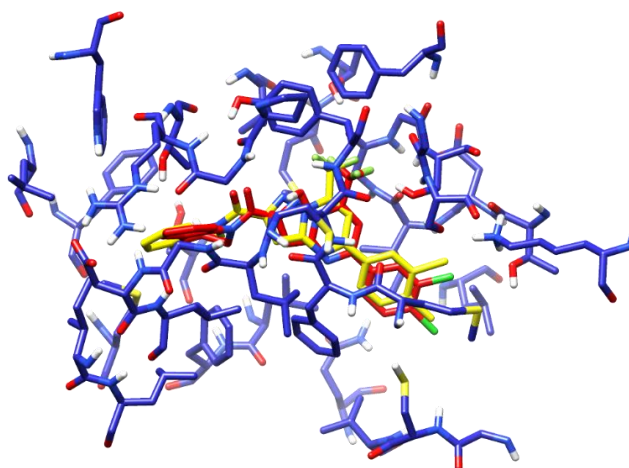


*Fig. 26 Conformazione 1 (in verde), sovrapposta alla pose di input della dinamica (in rosso)*



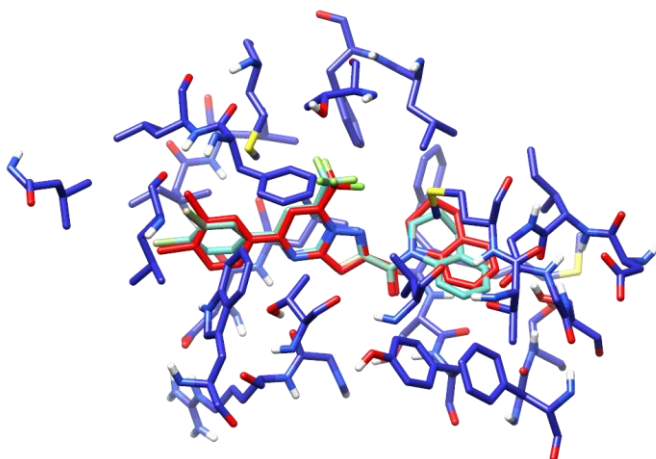
*Fig. 27 Conformazione 2 (in verde), sovrapposta alla pose di input dinamica (in rosso)*

- Fig. 23: l'RMSD durante la dinamica assume un andamento piuttosto lineare e la conformazione della pose rimane quasi uguale al minimizzato di input (Fig. 28);



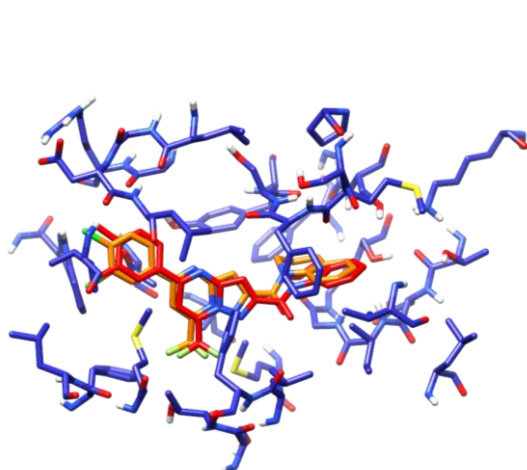
*Fig. 28 Average del risultato della dinamica (in giallo), sovrapposta alla pose di input (rosso)*

- Fig 24: in questo caso l'alto valore di RMSD è il risultato di una deviazione traslazionale della pose della dinamica rispetto a quello di partenza, dopo circa 40 step però la pose torna quasi a sovrapporsi a quello di partenza (Fig. 29), dando come risultato un andamento piuttosto stabile

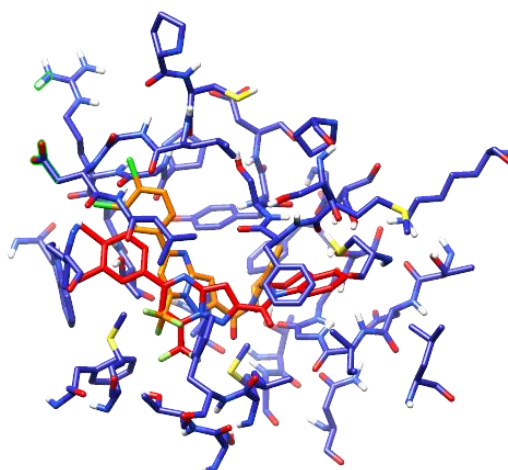


*Fig. 29 Pose del risultato della dinamica (in ciano), sovrapposta alla pose di input (rosso)*

- Fig. 25: come si vede dal grafico RMSD i primi 35 step della dinamica sono caratterizzati da una pose quasi identica a quella di input (Fig. 30), mentre i successivi step portano ad avere una pose strutturalmente diversa e piuttosto traslata in termini di coordinate (Fig. 31). Come descritto per la Fig. 22, questa nuova posizione, mantenuta per ben 14ns, risulta più stabile e favorita rispetto alla prima conformazione.



*Fig. 30 Pose 1 (in verde), sovrapposta alla pose di input della dinamica (in rosso)*



*Fig. 31 Pose 2 (in verde), sovrapposta alla pose di input della dinamica (in rosso)*

Dall'analisi di questi grafici si evince come il sistema tenda sempre a stabilizzarsi dopo circa 80 step di dinamica molecolare. È risultato perciò opportuno fare nuovi calcoli di valutazione di GBSA, PBSA, RMSD medio e % legame a idrogeno, considerando solo gli ultimi 100 step. Purtroppo, come dimostrato in tabella 3, questa analisi non ha portato ad un'individuazione marcata di una singola dinamica:

| Pose | GB 100<br>Kcal/mol | PB 100<br>Kcal/mol | RMSD<br>Lig<br>100 Å | RMSD<br>prot<br>100 Å | Stabilità<br>% | Leg H<br>%                    | Cluster | Pop      |
|------|--------------------|--------------------|----------------------|-----------------------|----------------|-------------------------------|---------|----------|
| 2    | -50.40             | -40.40             | 4.130                | 1.040                 | 60.00          | 0.500%                        | 1/5     | 991/1000 |
| 3    | -51.80             | -29.80             | 1.020                | 1.020                 | 100.0          | 50.00%<br>(F22/F23<br>SER209) | 1/5     | 161/1000 |
| 11   | -48.50             | -36.90             | 1.250                | 1.050                 | 85.00          | 0.600%                        | 3/5     | 3/1000   |
| 14   | -52.30             | -37.40             | 3.960                | 0.970                 | 85.00          | 0.000%                        | 3/5     | 183/1000 |

Tab. 3 Migliori risultati ottenuti dall'analisi della dinamica molecolare

### 3.5 Creazione di un Average e Studio di Self-Consensus-Docking

Dai dati forniti dalle valutazioni delle dinamiche molecolari effettuate, è apparso quindi evidente che non poteva essere intrapresa un'unica strada e che ulteriori approfondimenti dovevano essere svolti.

L'operazione finale, per decidere quale pose, quindi quale conformazione spaziale potesse assumere la molecola VS8 all'interno del sito catalitico della FAAH, è stata quella di creare un AVERAGE (vedi paragrafo 2.3), per le 4 migliori dinamiche finali, e su questo, fare uno studio di *Self-Consensus-Docking* per rilevare quale pose desse consenso migliore.

Le basi teoriche del Self-Consensus-Docking sono le stesse riportate al paragrafo 3.2, solo che in questo caso, i soggetti presi in causa sono:

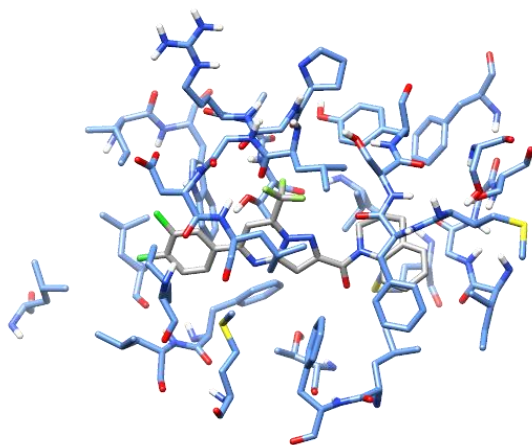
- *Docking della nuova conformazione del composto VS8 individuata dalla dinamica x, sulla proteina ricavata dalla stessa dinamica x.* Nel caso in cui la pose individuata dal docking sia la stessa di quella individuata dalla dinamica molecolare, si passa allo step successivo
- *Docking della pose VS8 di partenza, sulla proteina ricavata dalla dinamica x.* Esattamente come un calcolo di consensus docking, il parametro di valutazione viene

intrapreso in base a quanti programmi di docking portano al riassetamento spaziale della pose VS8, in modo da essere sovrapponibile alla pose individuata dalla dinamica molecolare, con un RMSD < 2.

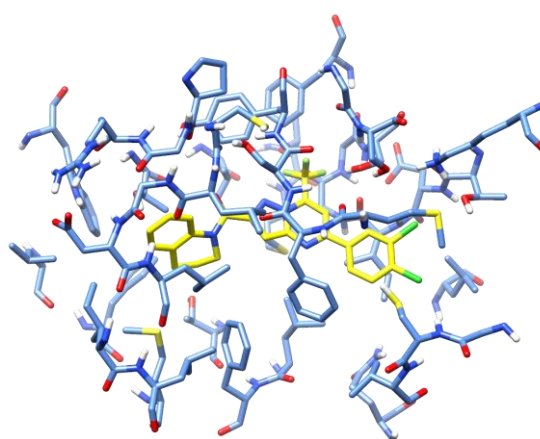
Dato che lo step 1 è stato passato con successo solo dalle pose 2,3 e 11 l'analisi che ha determinato l'individuazione di un binding mode esplicito per la FAAH, è stata decretata dal secondo step. Di seguito vengono riportati i risultati:

| N° Pose | N° Consensus |
|---------|--------------|
| 3       | 10/10        |
| 11      | 6/10         |
| 2       | 3/10         |

La pose 3 ha avuto miglior successo, con un Consensus massimo e come si può vedere in figura 33, rispetto alla pose VS8 (Fig 32) di partenza denota una diversa posizione di legame, più interna nel sito attivo, completamente capovolta sull'asse delle x e con il diclorofenile al limite sinistro del citosolic port, mentre il fluoro 22 sembra essere richiamato da una leggera interazione con l'OH della Ser241 catalitica. Sarà quindi presa come riferimento per studi di lead optimization mirati allo sviluppo di analoghi con maggiore attività inibitoria.



*Fig. 32 Pose VS8 all'interno del sito catalitico della FAAH*



*Fig. 33 Pose 3 all'interno della FAAH rinvenuta dagli studi di dinamica*



## CAPITOLO 4 CONCLUSIONI

La FAAH (*fatty acid amide hydrolase*) è il principale enzima deputato alla degradazione dell'anandamide, ligando endogeno dei recettori CB1 e CB2, e di altre etanolammidi bioattive. Essa è considerata un importante obiettivo farmacologico, soprattutto in quelle patologie che possono beneficiare di un aumento del tono endocannabinoide. La sua inattivazione ha dimostrato effetti terapeutici in numerosi studi preclinici, con potenziali applicazioni nel trattamento dell'infiammazione, del dolore neuropatico, delle malattie neurodegenerative, della depressione e dei disturbi metabolici e cardiovascolari.

L'utilizzo di inibitori selettivi verso questo enzima è emerso come un approccio in grado di aumentare il tono endocannabinoide in maniera evento-specifica, riducendo allo stesso tempo gli effetti psicotropi dati dall'uso prolungato di agonisti diretti dei recettori cannabinoidi. Numerose serie di inibitori sono note ad oggi in letteratura, quasi tutte caratterizzate dalla presenza di un carbonio elettrofilo in grado di legarsi covalentemente alla serina catalitica dell'enzima, in modo reversibile o irreversibile. La maggior parte di esse mostra selettività per la FAAH nei tessuti cerebrali, ma possiede target aggiuntivi in tessuti periferici come fegato e rene, tra cui molti isoenzimi coinvolti nel metabolismo di farmaci e profarmaci. La scoperta e lo sviluppo di inibitori prettamente non covalenti si pone come una valida strategia per cercare di migliorare il profilo di selettività e sicurezza mostrato dalle altre classi di inibitori, ed è stato perciò l'obiettivo del nostro studio.

La ricerca bibliografica ha portato alla luce un numero abbastanza esiguo di inibitori non covalenti. Attualmente sono disponibili soltanto due strutture cristallografiche della FAAH complessata con un inibitore di questo tipo: l'inibizione è ottenuta primariamente per complementarità di forma con il sito attivo e attraverso numerose interazioni idrofobiche, mentre non sono presenti legami ad idrogeno con determinati residui amminoacidici.

La prima parte del progetto è stata dedicata all'elaborazione di una procedura robusta di docking molecolare con i programmi Dock, Glide, Gold, Autodock, su un inibitore non-covalente (VS8) individuato tramite studi precedenti effettuati nel nostro laboratorio. Lo scopo di questa procedura è stato quello di individuare il reale binding mode del VS8 all'interno del sito catalitico della FAAH.

I risultati sono stati elaborati “clusterizzando” le migliori pose di ogni programma e prendendo in esame i rappresentanti di ogni singolo gruppo, i quali sono stati confrontati sulla base di energia di legame e di popolazione del cluster.

Solo 15 tra le pose analizzate sono state selezionate ed è stato realizzato uno studio di dinamica molecolare grazie al software AMBER14, in cui viene fornita energia e dunque libertà di movimento ai complessi proteina-ligando. Ciò permette di valutare la stabilità e di prevedere se il ligando ha mantenuto o meno le interazioni chiave con il sito catalitico della proteina. Se queste non sono sufficientemente forti, il complesso potrà andare incontro a dissociazione.

I risultati ottenuti dalla dinamica molecolare sono stati valutati attraverso l’analisi dei valori energetici di Generalized Born Surface Area (GBSA) e Poisson Boltzmann Surface Area (PBSA), dei valori di Root Mean Square Deviation (RMSD) e percentuale di legame a idrogeno.

Queste analisi hanno portato alla luce 4 possibili binding mode della pose VS8, sui quali è stata effettuata una procedura di Self-Consensus-Docking. Questo tipo di calcolo si affaccia alla tecnica del consensus docking, portata avanti nel nostro laboratorio, dove più alto è il numero di procedure di docking che predicono una determinata pose, più tale pose si avvicina a quella reale. Con il termine Self si indica che i soggetti non sono costituiti da nuove pose, ma sono costituiti dalle diverse conformazioni spaziali della stessa pose (VS8).

Alla fine di questo studio la pose che ha avuto miglior successo, con un Consensus massimo, sarà quindi presa come riferimento per studi di lead optimization mirati allo sviluppo di analoghi del composto VS8 con maggiore attività inibitoria

## ***BIBLIOGRAFIA***

- [1] X. HOLLISTER, L.E. PHARMAC. REV. 38, 1-20 (1986)
- [2] L. A.MATSUDA, S. J.LOLAIT, M. J.BROWNSTEIN , A. C.Y OUNG, T. I.BONNER, NATURE 1990, 346, 561.
- [3] Y.GAONI, R. MECHOULAM, J. AM. CHEM. SOC. 1964, 86, 1646.
- [4] W. A. DEVANE, L.HANUS , A.BREUER , R. G.PERTWEE , L. A.STEVENSON, G.GRIFFIN, D.GIBSON, A.MANDELBAUM, A.ETINGER , R.MECHOULAM, SCIENCE 1992, 258, 1946.
- [5] LABAR, G. AND MICHAUX, C. FATTY ACID AMIDE HYDROLASE: FROM CHARACTERIZATION TO THERAPEUTICS. CHEM. BIODIVERS. 2007;4:1882–1902
- [6] R. MECHOULAM,S. BEN SHABAT, L. HANUS, M. LIGUMSKY, N. E. KAMINSKI, A. R. SCHATZ, A. GOPHER, S. ALMOG, B. R. MARTIN, D. R. COMPTON, BIOCHEM. PHARMACOL. 1995, 50, 83.
- [7] T. SUGIURA, S. KONDO, A. SUKAGAWA, S. NAKANE, A. SHINODA, K. ITOH, A. YAMASHITA, K. WAKU, BIOCHEM. BIOPHYS. RES. COMMUN. 1995, 215, 89.
- [8] N.STELLA, P.SCHWEITZER , D. PIOMELLI, NATURE 1997, 388, 773.
- [9] MUCCIOLI GG. ENDOCANNABINOID BIOSYNTHESIS AND INACTIVATION, FROM SIMPLE TO COMPLEX. DRUG DISCOV TODAY 2010;15:474-83.
- [10] KOZAK KR, MARNETT LJ. OXIDATIVE METABOLISM OF ENDOCANNABINOID. PROSTAGLANDINS, LEUKOTRIENES ESSENT. FATTY ACIDS 2002;66:211-20.
- [11] K.TSUBOI, Y. X.SUN, Y.OKAMOTO , N.ARAKI, T.T ONAI, N.UEDA, J. BIOL. CHEM. 2005, 280, 11082.
- [12] B. Q.WEI, T. S.MIKKELSEN, M. K.MCKINNEY , E. S. LANDER, B. F.CRAVATT, J. BIOL. CHEM. 2006, 281, 36569.

- [13] BLANKMAN JL, SIMON GM, CRAVATT BF. A COMPREHENSIVE PROFILE OF BRAIN ENZYMES THAT HYDROLYZE THE ENDOCANNABINOIDE 2-ARACHIDONOYLGLYCEROL. CHEM BIOL 2007;14:1347-56.
- [14] CHICCA A, MARAZZI J, NICOLUSSI S, GERTSCH J. EVIDENCE FOR BIDIRECTIONAL ENDOCANNABINOIDE TRANSPORT ACROSS CELL MEMBRANES. J BIOL CHEM 2012;287:34660-82.
- [15] T. BISOGNO, M. MACCARONE. LATEST ADVANCES IN THE DISCOVERY OF FATTY ACID AMIDE HYDROLASE INHIBITORS. EXPERT OPIN. DRUG. DISCOV. 2013. H.LICHTMAN, C. C.SHELTON, T.ADVANI, B. F.CRAVATT, PAIN 2004, 109, 319.
- [16] S.PATEL, E. J.CARRIER , W.-S. V.HO , D. J.RADEMA CHER, S.CUNNINGHAM, D. S.REDDY , J. R. FALCK, B. F. CRAVATT, C.J . HILLARD, J. LIPID RES. 2005, 46, 342.
- [17] S. KATHURIA, S.GAETANI, D.F EGLEY, F.VALINO, A.DURANTI, A.TONTINI, M.MOR, G.T ARZIA, G.LA RANA, A.CALIGNANO , A.GIUSTINO , M.T ATTOLI, M. PALMERY, V.CUOMO , D.PIOMELLI, NAT. MED. 2003, 9, 76.
- [18] BACHUR, N.R., UDENFRIEND, S., 1966. MICROSOMAL SYNTHESIS OF FATTY ACID AMIDES. J. BIOL. CHEM. 241, 1308–1313.
- [19] SCHMID, P.C., ZUZARTE-AUGUSTIN, M.L., SCHMID, H.H.O., 1985. PROPERTIES OF RAT LIVER N-ACYLETHANOLAMINE AMIDOHYDROLASE. J. BIOL. CHEM. 260, 14145–14149
- [20] CRAVATT, B.F., GIANG, D.K., MAYFIELD, S.P., BOGER, D.L., LERNER, R.A., GILULA, N.B., MOLECULAR CHARACTERIZATION OF AN ENZYME THAT DEGRADES NEUROMODULATORY FATTYACID AMIDES. NATURE 1996;384, 83–87.
- [21] MILENI M, JOHNSON DS, WANG Z, ET AL. STRUCTURE-GUIDED INHIBITOR DESIGN FOR HUMAN FAAH BY INTERSPECIES ACTIVE SITE CONVERSION. PROC NATL ACAD SCI USA 2008;105:12820-4.
- [22] BRACEY MH, HANSON MA, MASUDA KR,ET AL. STRUCTURAL ADAPTATIONS IN AMEMBRANE ENZYME THAT TERMINATESENDOCANNABINOID SIGNALING. SCIENCE 2002;298:1793-6.

- [23] D.K. GIANG, B.F. CRAVATT, PROC. NATL. ACAD. SCI. U.S.A. 1997, 94, 2238
- [24] MIN X, THIBAUT ST, PORTER AC, GUSTIN DJ, CARLSON TJ, XU H, LINDSTROM M, XU G, UYEDA C, MA Z, LI Y, KAYSER F, WALKER NP, WANG Z. DISCOVERY AND MOLECULAR BASIS OF POTENT NONCOVALENT INHIBITORS OF FATTY ACID AMIDE HYDROLASE (FAAH). PROC NATL ACAD SCI U S A. 2011;108(18):7379-84.
- [25] M. P. PATRICELLI, B. F. CRAVATT, BIOCHEMISTRY 38, 14125 (1999)
- [26] M. P. PATRICELLI, B. F. CRAVATT, J. BIOL. CHEM. 2000, 275, 19177
- [27] M. K. MCKINNEY, B. F. CRAVATT, J. BIOL. CHEM. 2003, 278, 37393
- [28] DONOVAN 1845, BERRY E MECHOULAM, 2002
- [29] WILLIAMS E KIRKHAM, 1999;. HAO ET AL, 2000
- [30] JAMSHIDI E TAYLOR, 2001
- [31] WILLIAMS E KIRKHAM, 2002
- [32] DI MARZO ET AL., 2001B
- [33] BLU'HER, M. ET AL. DYSREGULATION OF THE PERIPHERAL AND ADIPOSE TISSUE ENDOCANNABINOID SYSTEM IN HUMAN ABDOMINAL OBESITY. DIABETES. 2006;55, 3053–3060
- [34] MATIAS, I. ET AL. REGULATION, FUNCTION, AND DYSREGULATION OF ENDOCANNABINOIDS IN MODELS OF ADIPOSE AND BETA-PANCREATIC CELLS AND IN OBESITY AND HYPERGLYCEMIA. J. CLIN. ENDOCRINOL. METAB. 2006;91, 3171–3180.
- [35] DI MARZO, V. ET AL. (2008) THE ROLE OF ENDOCANNABINOIDS IN THE REGULATION OF GASTRIC EMPTYING: ALTERATIONS IN MICE FED A HIGH-FAT DIET. BR. J. PHARMACOL. 2008;153, 1272–1280.
- [36] REVIEWD IN MECHOULAM, 1986; IVERSEN, 2000; MECHOULAM E HANUS, 2000

- [37] SOFIA ET AL, 1973;. FORMUKONG ET AL, 1988;. CALIGNANO ET AL., 1998;. COSTA ET AL, 2004A, B; GUINDON ET AL., 2006; ULUGOL ET AL, 2006
- [38] SOFIA ET AL, 1973;... MARTIN ET AL, 1996; SMITH ET AL, 1998;. GUINDON E BEAULIEU, 2006
- [39] BUXBAUM, 1972;. BLOOMETAL, 1977; ICHTMANANDMARTIN, 1991A, B; FRIDE E MECHOULAM, 1993; GUINDON E BEAULIEU, 2006
- [40] HERZBERG ET AL, 1997;.. PONTI ET AL, 2001; FOX ET AL, 2001;. GUINDON E BEAULIEU, 2006
- [41] TSOU ET AL, 1996; RICHARDSONETAL, 1998A, B, C;. LIETAL, 1999;. MARTINETAL, 1999B;. GUINDON ET AL, 2006
- [42] GUINDON E BEAULIEU, 2006; GUINDON ET AL, 2006;.. ULUGOL ET AL, 2006
- [43] MARTIN ET AL, 1993;. HOHMANN ET AL, 1995, 1998, 1999;.. MARTIN ET AL, 1995, 1996, 1998, 1999A, B; RICHARDSON ET AL, 1997, 1998A, B;. MENG ET AL, 1998;. STRANGMAN ET AL, 1998;. HOHMANN E WALKER, 1999;. FOX ET AL, 2001
- [44] YAKSH, 1981; LICHTMAN E MARTIN, 1991A , B, WELCH E STEVENS, 1992;. RICHARDSON ET AL, 1997, 1998A, B; HOHMANN ET AL, 1998;. CHAPMAN, 1999;. DREW ET AL, 2000;. NADERI ET AL., 2005; SUPLITA ET AL, 2006
- [45] CALIGNANO ET AL, 1998;. RICHARDSON ET AL, 1998C;. HOHMANN E HERKENHAM, 1999; FOX ET AL, 2001;. JOHANEK ET AL, 2001;.. KELLY ET AL, 2003; JOHANEK E SIMONE, 2004. JORDT ET AL, 2004; AMAYA ET AL, 2006
- [46] HERKENHAM ET AL, 1990, 1991 A;. HOHMANN E HERKENHAM, 1998, 1999; HOHMANN ET AL, 1999;. SANUDO-PENA ET AL, 1999°
- [47] CALIGNANO ET AL, 1998;. HANUS ET AL, 1999;. MALAN ET AL, 2001;.. CLAYTON ET AL, 2002; IBRAHIM ET AL, 2003, 2005;.. NACKLEY ET AL, 2003A, B, 2004; QUARTILHO ET AL, 2003;.. ELMES ET AL, 2004; HOHMANN E COLL ., 2004; SCOTT ET AL, 2004;.. WHITESIDE ET AL., 2005; IBRAHIM ET AL, 2006

- [48] REVIEWD A WARE E BEAULIEU, 2005, 2006
- [49] REVIEWD IN KLEIN ET AL, 1998, 2003;.. WALTER E STELLA, 2004; KLEIN, 2005
- [50] REVIEWD A KLEIN ET AL, 1998, 2003;.. WALTER E STELLA, 2004; KLEIN, 2005
- [51] REVIEWD IN VETRO, 2001; MECHOULAM ET AL, 2002A, B;.. GRUNDY, 2002; CROXFORD, 2003; DRYSDALE E PLATT, 2003; JACKSON ET AL, 2005A;.. RAMOS ET AL, 2005
- [52] MOLINA-HOLGADO ET AL, 1997;.. MARSICANO E LUTZ, 1999; GERDEMAN ET AL., 2002; RIVISTO IN ALGERI, 2002 ; ROBBE ET AL 2002;.. AZADETAL, 2003;.. FREUND ET AL, 2003; GERDEMAN E LOVINGER, 2003; PIOMELLI, 2003;.. MATO ET AL, 2004
- [53] WATZL ET AL, 1991;.. ZHENG ET AL, 1992;.. FISCHER STENGER ET AL, 1993;.. CABRAL E FISCHER-STENGER, 1994;.. KUSHER ET AL, 1994; BURNETTE-CURLEY E CABRAL, 1995; CABRAL ET AL, 1995;.. REVIEWD DI FRIEDMAN ET AL, 1995;.. ZHENG E SPECTER, 1996; SHOHAMI ET AL, 1997;.. NEWTON ET AL, 1998;.. SRIVASTAVA ET AL, 1998;.. GALLILY ET AL, 2000;.. KLEIN ET AL, 2000A, B, 2003. SMITH ET AL, 2000;.. CARLISLE ET AL, 2002; GERMAIN ET AL, 2002;.. KILLESTEIN ET AL, 2003;.. KAPLAN ET AL., 2005; RAMIREZ ET AL., 2005; RIVISTO IN FRIEDMAN ET AL, 1995;.. STELLA 2004; WALTER E STELLA, 2004; CORREA ET AL., 2005; CROXFORD E YAMAMURA 2005; KLEIN, 2005
- [54] GRIGORENKO ET AL., 2002
- [55] MOLINA-HOLGADO ET AL., 2002
- [56] KIM ET AL., 2005
- [57] KHASIKOV ET AL., 2004
- [58] CAULFIELD E BROWN, 1992; MACKIE E HILLE, 1992; MACKIE ET AL, 1993;.. NADLER ET AL, 1995;.. VENANCE ET AL, 1995;.. SHOHAMI ET AL, 1997;.. HAMPSON ET AL, 2000B;.. OZETAL 2000 2004; CHEMIN ET AL, 2001;.. MAINGRET ET AL, 2001;.. NOGUERON ET AL, 2001;.. ROBBE ET AL, 2001;.. WILSON E NICOLL, 2001; WILSON ET AL, 2001;.. NICHOLSON ET AL, 2003. , GUO E IKEDA, 2004. DELCARMEN ET AL, 2005; DELCARMEN GODINO ET AL., 2005;.. ZHUANG ET AL, 2005

- [59] ESHHAR ET AL, 1995;. HAMPSON ET AL, 1998;. CHENANDBUCK, 2000;  
REVIEWDEDINHAMPSON ET AL, 2000A;.. MARSICANO ET AL, 2002A
- [60] LEKER ET AL., 2003
- [61] HOLM ET AL., 2005, FADEN, 2002; MAAS ET AL., 2004
- [62] BAYIRETAL., 2003
- [63] SHOHAMI ET AL., 1993, 1995
- [64] SOSPEDRA E MARTIN, 2005
- [65] GRUNDY, 2002; PERTWEE, 2002; BAKER E PRYCE, 2003; CROXFORD, 2003;  
KILLESTEIN ET AL, 2004;. SIRVEN E BERG, 2004; JACKSON ET AL, 2005A;. PRYCE E BAKER,  
2005 , ROBSON, 2005; SMITH, 2005
- [66] HERKENHAM ET AL, 1990, 1991A, B;. MAILLEUX E VANDERHAEGHEN, 1992; TSOU ET  
AL, 1998;. HOHMANN E HERKENHAM, 2000 MOLDRICH E WENGER, 2000; HOWLETT ET AL,  
2002
- [67] BISOGNO ET AL, 1999A;.. DI MARZO ET AL, 2000
- [68] CRAWLEY ET AL, 1993;. FRIDE E MECHOULAM, 1993; WICKENS E PERTWEE, 1993;  
SMITH ET AL, 1994;. ROMERO ET . AL, 1995A, B, 2002B;. RIESAMINATI SAN~UDO-PEN~A ET  
AL, 1999B
- [69] ZENG ET AL, 1999;. PAGINA ET AL, 2000;. ROMERO ET AL, 2000;.. LASTRES-BECKER  
ET AL, 2001A, B, 2002A , B;. GONZALEZ ET AL, 2006
- [70] GLASSET AL, 1993, 2000, 2004;. LASTRES-BECKER ET AL, 2001A;.. RIVEDUTI ROMERO  
ET AL, 2002B
- [71] GLASSETAL, 1997A;.. MILLERETAL, 1998; SAN~UDO-PEN~AAND WALKER, 1998;  
GIUFFRIDA ET AL, 1999;. RODRIGUEZ DE FONSECA ET AL, 2001;. BROTHIE, 2003; VAN DER  
STELT E DI MARZO, 2003;. DE LAGO ET AL, 2004°
- [72] DODICI ET AL., 2003



- [73] REVIEWD IN BROTHIE, 2003
- [74] GERDEMAN E LOVINGER, 2001; GERDEMAN ET AL, 2002;.. GUBELLINI ET AL, 2002
- [75] SANUDO-PENA ET AL, 1998, 1999B
- [76] ANDERSON ET AL, 1995;.. MANEUF ET AL, 1997;... SAN~UDO-PEN~A ET AL, 1998
- [77] LASTRES-BECKER ET AL, 2005
- [78] CONSROE, 1998; MULLER-VAHL ET AL, 1999A;.. ROMERO ET AL, 2002;.. BROTHIE, 2003; CROXFORD , 2003; CROXFORD E MILLER, 2003
- [79] REVIEWD IN PIOMELLI, 2003; LUTZ, 2004
- [80] MARSICANO E LUTZ, 1999; ALGER, 2002, 2004; GERDEMAN ET AL, 2002;.. ROBBE ET AL., 2002; AZAD ET AL, 2003;.. FREUND ET AL, 2003; GERDEMAN E LOVINGER, 2003; KIM E ALGER 2004; ISOKAWA E ALGER, 2005
- [81] GAONI E MECHOULAM 1964
- [82] REVIEWD GORDON E DEVINSKY 2001; LUTZ, 2004
- [83] WALLACE ET AL., 2003°
- [84] BAN, 2004
- [85] REVIEWD IN THAKER E CARPENTER, 2001; LEWIS ET AL, 2005
- [86] EMRICHETAL 1997, REVIEWDED IN UJIKE AND MORITA 2004;.. LAVIOLETTE E GRACE, 2006
- [87] LEWEKEETAL, 1999;.. DE MARCHI ET AL., 2003; GIUFFRIDA ET AL, 2004
- [88] REVIEWD IN ZUARDI ET AL., 2006
- [89] PACHER E KECSKEMETI, 2004

- [90] PATON E PERTWEE, 1973; ASHTON ET AL, 1981, 2005; VIVEROS ET AL, 2005
- [91] HALLER ET AL., 2004A, B
- [92] REVIEWD IN VIVEROS ET AL., 2005
- [93] HERKENHAM ET AL., 1990, 1991A , B; VETRO ET AL, 1997B;.. KATONA ET AL, 2001; HA'JOS E FREUND, 2002; TZAVARA ET AL, 2003;.. PISTIS ET AL, 2004
- [94] HILLETAL., 2006
- [95] GOBBI ET AL., 2005Ù
- [96] FREEMAN, 1972, 1982; PIVIK ET AL, 1972;.. BARRATT ET AL, 1974;.. FEINBERG ET AL, 1975, 1976
- [97] MONTI 1977;.. BUONAMICI ET AL, 1982
- [98] MURILLO-RODRIGUEZ ET AL., 1998, 2001
- [99] BOGER ET AL., 1998°
- [100] REVIEWD DA MARTIN E WILEY, 2004; AAPRO, 2005; HALL ET AL., 2005
- [101] DARMANI, 2001
- [102] PRKERETAL., 2004
- [103] GEBREMEDHIN TAAL, 1999;.. JA'RAHI ET AL, 1999;.. WAGNER ET AL, 2001B;.. WAGNER ET AL, 2005
- [104] BONZ ET AL, 2003;.. MASLOV ET AL, 2005
- [105] BONZ ET AL, 2003;.. BA'TKAI ET AL, 2004B;.. PACHER ET AL, 2004B, 2005A, B, D; ENGELI ET AL., 2005; WAGNER ET AL, 2005
- [106] GEBREMEDHIN ET AL., 1999
- [107] VARGA ET AL, 1995;.. LAGO ET AL, 1997B;.. PACHER ET AL, 2004B, 2005D

- [108] MALINOWSKA ET AL., 2001
- [109] PACHER ET AL, 2004
- [110] VARGA ET AL., 1995, 1996; CALIGNANO ET AL., 1997; PACHER ET AL., 2004B, 205A,D
- [111] LAKEETAL, 1997°
- [112] LAKEETAL, 1997°
- [113] WAGNER ET AL, 1997;. CAIAZZO ET AL., 2002
- [114] VARGA ET AL, 1998;. WANG ET AL, 2001;.. LIU ED ALTRI, 2003A;. BA'TKAI ET AL, 2004A; GARDINER ET AL., 2005; KADOI ET AL. , 2005
- [115] WAGNER ET AL., 2001A, 2003
- [116] MATSUDA ET AL., 2005
- [117] WAGNER ET AL, 1997, 2001A, B;. VARGA ET AL., 1998
- [118] REVIEWD IN CSISZAR ET AL., 2005; LIBBY E THEROUX, 2005
- [119] ROTH, 2005; STEFFENS ET AL., 2005
- [120] REVIEWD LEMBERGER, 1980;. TASHKIN ET AL, 2002
- [121] TASHKIN ET AL, 1973;. VACHON ET AL., 1973
- [122] REVIEWD IN WALSH ET AL, 2003;. SALA ET AL., 2005
- [123] WALSH ET AL, 2003;. HALL ET AL., 2005
- [124] REVIEWD IN PAROLARO ET AL, 2002;. GUZMA'N ET AL, 2002;. GUZMA'N, 2003; JONES E HOWL, 2003; VELASCO ET AL., 2004; PATSOS ET AL., 2005

[125] REVIEWED IN BIFULCO AND DiMARZO, 2002; PAROLARO ET AL, 2002;. GUZMAN ET AL, 2002;. GUZMAN, 2003; JONES AND HOWL, 2003;. VELASCO ET AL, 2004;. PATSOS ET AL, 2005

[126] DE PETROCELLIS ET AL, 1998;. SANCHEZ ET AL, 1998, 2003;. JACOBSSON ET AL, 2000;. MACCARRONE ET AL, 2000B;. SARKER ET AL, 2000;. MCKALLIP ET AL, 2002A, B; FOWLER ET AL., 2003; JONSSON ET AL, 2003;. MIMEAULT ET AL, 2003; BIFULCO ET AL., 2004; CONTASSOT ET AL, 2004A, B.; HINZ ET AL., 2004; JOSEPH ET AL., 2004; KOGAN ET AL., 2004; MASSI ET AL., 2004; NITHIPATIKOM ET AL., 2004; ALLISTER ET AL., 2005; ELLERT MIKLASZEWSKA ET AL., 2005; HERRERA ET AL, 2005, 2006.; LOMBARD ET AL., 2005; POWLES ET AL., 2005; SARFARAZ ET AL., 2005; VACCANI ET AL., 2005; CARRACEDO ET AL., 2006; GRIMALDI ET AL., 2006; LIGRESTI ET AL., 2006B

[127] EDWARDS PD, MEYER EF, VIJAYALAKSHMI J, TUTHILL PA, ANDISIK DA, GOMES B, ET AL. DESIGN, SYNTHESIS, AND KINETIC EVALUATION OF A UNIQUE CLASS OF ELASTASE INHIBITORS, THE PEPTIDYL  $\alpha$ -KETOBENZOXAZOLES, AND THE X-RAY CRYSTAL STRUCTURE OF THE COVALENT COMPLEX BETWEEN PORCINE PANCREATIC ELASTASE AND AC-ALA-PRO-VAL<sup>2</sup>-BENZOXAZOLE. J AM CHEM SOC 1992; 114(5): 1854-63

[128] EDWARDS PD, WOLANIN DJ, ANDISIK DW, DAVIS MW. PEPTIDYL  $\alpha$ -KETO HETEROCYCLIC INHIBITORS OF HUMAN NEUTROPHIL ELASTASE. 2. EFFECT OF VARYING THE HETEROCYCLIC RING ON IN VITRO POTENCY. J MED CHEM 1995; 38(1): 76-85

[129] MILENI M, KAMTEKAR S, WOOD DC, BENSON TE, CRAVATT BF, STEVENS RC. CRYSTAL STRUCTURE OF FATTY ACID AMIDE HYDROLASE BOUND TO THE CARBAMATE INHIBITOR URB597: DISCOVERY OF A DEACYLATING WATER MOLECULE AND INSIGHT INTO ENZYME INACTIVATION. J MOL BIOL 2010; 400(4): 743-54

[130] DASSE, O., PUTMAN, D., COMPTON, T. R., PARROTT, J. A. METABOLICALLY STABILIZED INHIBITORS OF FATTY ACID AMIDE HYDROLASE, AND USE IN THE TREATMENT OF PAIN. WO2008063714 (2008).

[131] APODACA, R., BREITENBUCHER, J.G., PATTABIRAMAN, K., SEIERSTAD, M., XIAO, W. PIPERAZINYL AND PIPERIDINYL UREAS AS MODULATORS OF FATTY ACID AMIDE HYDROLASE. WO2006074025 (2006).

- [132] YANG, W., GAO, X., WANG, B. BORONIC ACID COMPOUNDS AS POTENTIAL PHARMACEUTICAL AGENTS. *MED RES REV* 2003; 23(3): 346-68.
- [133] ADAMS, J., BEHNKE, M.L., CASTRO, A.C., EVANS, C.A., GRENIER, L., GROGAN, M.J., LIU, T., SNYDER, D.A., TIBBITTS, T.T. PREPARATION OF ARYLBORONATES AS INHIBITORS OF FATTY ACID AMIDE HYDROLASE. WO2008063300 (2008)
- [134] MINKKILA A, SAARIO SM, KA SNA NEN H, LEPPA NEN J, POSO A, NEVALAINEN T. DISCOVERY OF BORONIC ACIDS AS NOVEL AND POTENT INHIBITORS OF FATTY ACID AMIDE HYDROLASE. *J MED CHEM* 2008; 51(22): 7057-60.
- [135] ZHANG D, SARAF A, KOLASA T, ET AL. FATTY ACID AMIDE HYDROLASE INHIBITORS DISPLAY BROAD SELECTIVITY AND INHIBIT MULTIPLE CARBOXYLESTERASES AS OFF-TARGETS. *NEUROPHARMACOLOGY* 2007;52:1095-105.
- [136] GUSTIN DJ, MA Z, MIN X, ET AL. IDENTIFICATION OF POTENT, NONCOVALENT FATTY ACID AMIDE HYDROLASE (FAAH) INHIBITORS. *BIOORG MED CHEM LETT* 2011;21:2492-6.
- [137] WANG X, SARRIS K, KAGE K, ET AL. SYNTHESIS AND EVALUATION OF BENZOTHIAZOLE-BASED ANALOGUES AS NOVEL, POTENT, AND SELECTIVE FATTY ACID AMIDEHYDROLASE INHIBITORS. *J MED CHEM* 2009;52:170-80.
- [138] TIAN G, PASCHETTO KA, GHARAH DAGHI F, ET AL. MECHANISM OF INHIBITION OF FATTY ACID AMIDE HYDROLASE BY SULFONAMIDE-CONTAINING BENZOTHIAZOLES: LONG RESIDENCE TIME DERIVED FROM INCREASED KINETIC BARRIER AND NOT EXCLUSIVELY FROM THERMODYNAMIC POTENCY. *BIOCHEMISTRY* 2011;50:6867-78.
- [139] MUCCIOLI, G. G.; FAZIO, N.; SCRIBA, G. K. E.; POPPITZ, W.; CANNATA, F.; POUPART, J. H.; WOUTERS, J.; LAMBERT, D. M. J. *MED. CHEM.* 2006, 49, 417.
- [140] NAKABAYASHI, M.; REGAN, M. M.; LIFSEY, D.; KANTOFF, P. W.; TAPLIN, M.-E.; SARTOR, O.; OH, W. K. *BJU INT.* 2005, 96, 783–786
- [141] MICHAUX, C.; MUCCIOLI, G. G.; LAMBERT, D. M.; WOUTERS, J. *BIOORG. MED. CHEM. LETT.* 2006, 16, 4772.

- [142] CLAY W SCOTT , GAOCHAO TIAN , XIAO HONG YU, KATHY A. PASCHETTO , DEE E. WILKINS , LUC MEURY , CHANG QING CAO , JEFFREY VARNES , PHILIP D. EDWARDS. BIOCHEMICAL CHARACTERIZATION AND IN VITRO ACTIVITY OF AZ513, A NONCOVALENT, REVERSIBLE, AND NONCOMPETITIVE INHIBITOR OF FATTY ACID AMIDE HYDROLASE. EUROPEAN JOURNAL OF PHARMACOLOGY 2011;667:74–79.
- [143] JAYENDRA ZAVERBHAI PATEL, TEIJA PARKKARI, TUOMO LAITINEN, AGNIESZKA A KACZOR, SUSANNA M SAARIO, JUHA R SAVINAINEN, DINA NAVIA-PALDANIUS, MARIATERESA CIPRIANO, JUKKA LEPPÄNEN, IGOR O. KOSHEVOY, ANTTI POSO, CHRISTOPHER J. FOWLER, TAPIO NEVALAINEN, AND JARMO T LAITINEN. CHIRAL 1,3,4-OXADIAZOL-2-ONES AS HIGHLY SELECTIVE FAAH INHIBITORS. 2013; 56(21):8484-96.
- [144] KAUB, C.; GAWLUGARI, S.; KINCAID, J.; JOHNSON, R. J.; O'MAHONY, D. J. R.; ESTIARTE MARTINEZ, M.; DUNCTON, M., RENOVIS, WO 2010039186, 2010.
- [145] APODACA, R.; BREITENBUCHER, J. G.; CHAMBERS, A. L.; DENG, H.; KAWRYLUK, N. A.; KEITH, J. M.; MANI, N. S.; MERIT, J. E.; PIERCE, J. M.; SEIERSTAD, M.; XIAO, W., JANSSEN PHARMACEUTICALS, WO 2009105220, 2009
- [146] LIN, L. S.; CHIODA, M. D.; LIU, P.; NARGUND, R. P., MERCK, WO 2009152025, 2009
- [147] APPLICATION OF A FLAP-CONSENSUS DOCKING